EEG Signal Classification Using the Event-Related Coherence and Genetic Algorithm

Chunying Fang [1,2], Haifeng Li[1], and Lin Ma[1]

[1] Harbin Institute of Technology, School of Computer Science and Technology, Harbin, China
{Lihaifeng,malin_li}@hit.edu.cn
[2] Heilongjiang Institute of Science and Technology, School of Computer and Information Engineering, Harbin, China
fcy3333@163.com

Abstract. The reliable operation of brain-computer interfaces (BCIs) based on spontaneous electroencephalogram (EEG) signals requires an accurate classification and recognition of multichannel EEG. The design of EEG representations and classifiers for BCI are open research questions whose difficulty stems from the need to extract complex spatial and temporal patterns from noisy multidimensional time series obtained from EEG measurements. This paper proposes a Genetic algorithm (GA) and Support Vector Machine (SVM) hybrid approach to accomplish this EEG classification task for potential BCI applications. An Oddball stimulus program and evoked event-related coherence program were designed to evaluate our method. The present study systematically evaluates the performance of the one channel pair event-related coherence feature set for EEG signal classification of auditory task. A GA approach for feature selection is presented which used to reduce the dimension of event-related coherence feature parameters. With the base classifiers of SVM, classification experiments are carried out upon real EEG recordings. Experimental results suggest the feasibilities of the new feature set, and we also derive some valuable conclusions on the performance of the EEG signal classification methods. The high recognition rates and the method's procedural and computational simplicity make it a particularly promising method for achieving real-time BCI system based on evoked potential event-related coherence in the future.

Keywords: EEG signal classification, Event-related coherence, Genetic algorithm, BCI.

1 Introduction

The electroencephalography(EEG) classification is one important part of the brain computer interface (BCI)[1], EEG is relatively more convenient, harmless and inexpensive than other methods[2] which provides a direct measure of cortical activity with millisecond temporal resolution[3]. By training the computer to recognize and classify EEG signals, users could manipulate the machine by merely thinking about what they want it to do within a limited set of choices[4]. Particularly relevant to the present study is a growing number of EEG classification studies which depends on

D. Liu et al. (Eds.): BICS 2013, LNAI 7888, pp. 92–100, 2013.

both the features and the classification algorithm employed. A great variety of features have been attempted to design BCI such as amplitude values of EEG signals [5], Band Powers (BP) [6], Power Spectral Density (PSD) values [7] [8], Auto Regressive (AR) and Adaptive Auto Regressive (AAR) parameters [9] [10], Time-frequency features [11] and inverse model-based features [12] [13] [14]. The used classification algorithms divided into five different categories: linear classifiers, neural networks, nonlinear Bayesian classifiers, nearest neighbor classifiers and combinations of classifiers in BCI systems [15]. However, as we know it may be difficult to build a good single classifier if feature parameters are high dimensionality and the training set is comparatively small.

Finding a suitable representation of EEG signals is the key to learning a reliable discrimination [16, 17].Oscillatory states are the most remarkable features of EEG activity, according to this view, the rhythmic synchronization during oscillatory states can serve to enhance perception, learning, and the transmission of neuronal signals between different regions of the brain[18].The EEG coherence analysis gives important information on EEG changes in long distance connections in brain areas upon application of perception/cognitive stimulations[19]. Until now, there is no study in the literature related to the analysis of event-related coherence as the feature parameters in the EEG signals classification. In a number of experiments, we found that the event-related coherence (ERCoh) from two bipolar channels (F4-M2) over the frontal and temporal areas during auditory change could be significant differentiated at different times (p=0.035), primarily within low alpha (8-10Hz) frequency band. Based on experience with the auditory change study, we sought to replicate the design using an ERCoh-based BCI. Furthermore, we discriminate the standard and deviant auditory stimulus. This article combined feature selection technique with the aim of reducing the number of required trials. The classifier used is a Support Vector Machine [20]. The results show that, the classification accuracy of the proposed method reaches 93% as compared to the current reported best accuracy of 84%. To this end, firstly, we designed the Oddball stimulus program and evoked potential event-related coherence experimental program, the collected EEG signals is pre-processed and pattern recognition which contained evoked potential signal, the evoked potential event-related coherence BCI system is established based on signal acquisition and processing model. The high recognition rates and the method's procedural and computational simplicity make it a particularly promising method for achieving real-time BCI system based on evoked potential event-related coherence in the future.

2 Method and Experiments

This research attempts to find a new feature parameter set and optimum algorithm to deal with the EEG signal to achieve the increase in the accuracy. The proposed model in five phases (see Fig.1), preprocessing, feature extraction, feature selection, classification, with EEG signal finally is classified into standard auditory stimuli or deviant auditory stimuli.

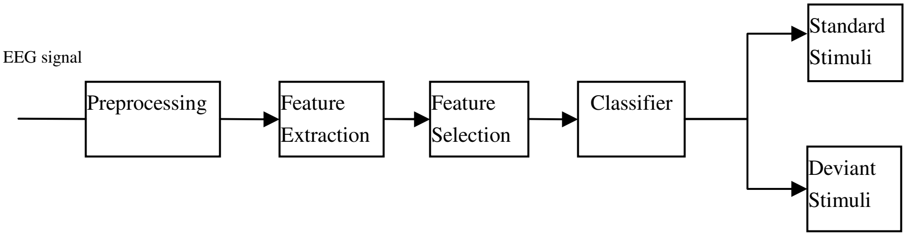

Fig. 1. Event-related coherence BCI system model

2.1 Preprocessing

The continuous electroencephalogram (EEG) was collected with Neuroscan NuAmps Amplifier, using Quick Cap with 64-channel Ag/AgCl electrodes according to the extended international 10-20 system. The reference electrode was placed on the nose tip. The vertical EOG was recorded from the right eye by supra- and infra-orbital electrodes, and horizontal EOG were recorded from electrodes on the outer canthi of both eyes. EEG and EOG signals were amplified from DC to 100 Hz at a sampling rate of 500 Hz. The electrode impedance was less than 5 kΩ throughout the experiment. After EOG artifact correction, the EEG was transformed offline to the epoch from 50ms pre-stimulus to 550ms post-stimulus. The trials contaminated with artifacts greater than ±100μV were rejected before averaging. The EEG segments were averaged separately for 150 ms with 50ms duration conditions, and the averaged ERPs were smoothed through a low-pass digital filter at 20 Hz (24 dB/octave).

2.2 Feature Extraction

We found that the event-related coherence (ERCoh) from two bipolar channels (F4-M2) over the frontal and temporal areas during auditory change could be significant differentiated at different times (p=0.035), primarily within low alpha (8-10Hz) frequency band. So we extracted the event-related coherence of F4-M2 during auditory change in the low alpha frequency band. Event-related coherence is a frequency dependent measure of the degree of linear relatedness between two channels. This symmetric measurement is computed from a collection of EEG epochs sampled from either ongoing or event-related activity. High coherence implies that amplitudes at a given frequency are correlated across EEG samples, moreover, that tends to be a constant phase angle (or time lag) between the two signals [21].

$$R_{xy} = \frac{\sum_i (x_i - \overline{x})(y_i - \overline{y})^*}{\sqrt{\sum_i (x_i - \overline{x})(x_i - \overline{x})^* \sum_i (y_i - \overline{y})(y_i - \overline{y})^*}} \qquad (1)$$

Where x_i and y_i is a pair of real numbers sampled on occasion *i*. Each time series can be translated to the frequency domain as a frequency spectrum of complex numbers $x_i(f)$ and $y_i(f)$. The result is a complex correlation spectrum; finally, the coherence spectrum consists of set of real numbers ranging between 0 and 1, with 0 in the case of independence and 1 in the case of a perfect linear relationship. For each frequency, this number measures the proportion of variance in the data that can be accounted for a best-fit linear relationship between the two variables. ERCoh is computed from epoch EEG data using the coherence formulas already given. However, the frequency of interest is preselected, and the results are a function of time with respect to the event at time zero. The real and imaginary parts come from sweep-by-sweep complex demodulation rather than from sweep-by -sweep FFT.

2.3 Feature Selection

Reducing the number of features will help the classifier learn a more robust solution and achieve a better generalization performance. Feature selection algorithms fall into two categories based on whether or not they perform feature selection independently of the learning algorithm that constructs the classifier. If the technique performs feature selection independently of the learning algorithm, it follows a filter approach. Otherwise, it follows a wrapper approach [22]. Genetic algorithms can find the most efficient features of the whole space, it has been demonstrated the most efficient feature selection method for learning areas and hence has less chance to get local optimal solution than other algorithms [23] [24] [25].

Genetic algorithm belongs to the wrapper approach; therefore, classifier is very important, and we use support vector machine (SVM) classifier in this paper. The algorithm flow is shown in Figure 2.

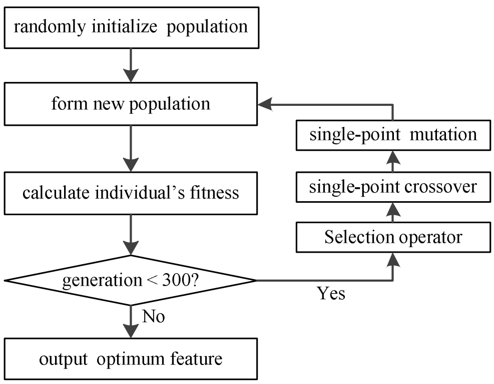

Fig. 2. The genetic algorithm flow

Encode Problem. Firstly, we must encode the problem. Here, we use a binary coded chromosome GA to select optimum feature descriptor subset for event-related coherence. Each code, namely a chromosome, corresponds to a solution of the problem.

Each gene on chromosome represents an input of the independent feature and its value can only be "0" or "1". If a gene's value is "1", it indicates that the corresponding feature is involved in the selected feature descriptor subset. On the contrary, the "0" indicates that the corresponding feature is not involved in the selected feature descriptor subset.

Initial Population. Let m as the number of feature descriptors, N the size of population. Commonly, population size N is $20 < N < 100$ and we use 60 in this paper. Chromosome of m genes is used to represent whether the corresponding feature is involved in the selected feature descriptor subset. In initial population $P = \{p_1, p_2, \cdots, p_N\}$, the genes of all individuals are randomly generated. Namely, each gene in a chromosome has value "0" or "1" randomly.

Fitness Function. Classification accuracy of SVM classifier is used to evaluate the fitness of individuals. In detail, we use the reciprocal of sum of the test set's squared errors as the fitness function, it is quite straightforward to see that.

$$f(p_k) = 1 \Big/ \sum_{i=1}^{n} (\hat{t}_i - t_i)^2 \tag{2}$$

Where $\hat{T} = \{\hat{t}_1, \hat{t}_2, \cdots, \hat{t}_n\}$ represent the predictive value of the test set classified by SVM using feature descriptor subset which p_k represents, $T = \{t_1, t_1, \cdots, t_n\}$ represent the true value of the test set, and n is the test data size.

Selection Operator. The selection operator determines an individual's genetic probability to the next generation population based on the individual's fitness. The processing is as follows:

Firstly, sum the fitness of all individual in the population:

$$F = \sum_{k=1}^{N} f(p_k) \tag{3}$$

Secondly, calculate the relative fitness of individual p_k in the population, which indicates the probability of the individual selected and inherited to the next generation:

$$\Pr(p_k) = f(p_k)/F \tag{4}$$

Finally, the simulated roulette to generate a random number between (0, 1), to determine the number of each individual selected. Obviously, larger individual's selection probability will lead to more repeatedly selected. Crossover operator uses single-point mode. The mutation operator in this paper uses the single-point mutation. We proceed with the next generation until the process reaches the maximum iteration 300 generation. When the process ends, the fittest individual is output as the optimum feature selection result.

2.4 Experiments

To this end, we designed an experiment which Fourteen healthy right-handed volunteers (4 males, 10 females; age=24.1±5.7 years) participated in this study. To eliminate circadian rhythm effects, the present experiment was carried out between 10:00 a.m. and 3:00 p.m. The task was an adaptation of the novelty "oddball" paradigm with an auditory modality, in which two types of stimuli, 1000 Hz frequent (75%) tones as non-target standard stimuli, infrequent (15%) 2000 Hz tones as targets, and rare (15%), All stimuli were presented binaurally at a sound level of 90 dB, with an exposure time of 100 ms and an inter-stimulus interval (ISI) of 600ms. The experimental session consisted of four blocks of 100 trials each, with a short time break between blocks. The EEG signal was recorded while participants were watching a self-selected, subtitled and silent film. Participants were instructed to ignore the acoustic stimuli. The continuous electroencephalogram (EEG) was collected with Neuroscan NuAmps Amplifier, using Quick Cap with 64-channel Ag/AgCl electrodes according to the extended international 10-20 system.

The event-related coherence was analyzed for low alpha (8-10 Hz) frequencies ranges, the 10 trials are were divided into one epoch in the time range of 50ms before and 550ms after the onset of auditory stimuli. The event-related coherence values for F4-M2 electrode pair were averaged across the single trials if the number of accepted trails is greater than three. The event-related coherence 302 dimension feature set is selected to form the new feature subset by genetic algorithm. The feature subset is examined with the classifier SVM, finally, the number of feature subset is 13 dimensions.

2.5 Results

SVM follows a procedure to find the separating hyperplane with the largest margin between two classes. It is based on statistical learning theory. The open source LibSVM is used to realize the SVM classifier. Firstly, to the single subject, we train the models using the two kinds(standard, deviant) of samples epoch which every epoch is consisted with 10 single trails and test is the same, the number of accept trails is greater than three to compute the event-related coherence, the accuracy rate is as shown in Table 1. Secondly, we choose ten subjects to analysis by the same method, the train set is 600 standard auditory stimuli and 100 deviant stimuli, and the test is 213 standard stimulus and 34 deviant stimulus group. The classified method is different. One is GMM(Gaussian mixture model) classifier, one is SVM(Support Vector Machine), the other is that the event-related coherence feature is selected and classified by GA and SVM, Results are as shown in table 2. Finally, the number of feature select subset is 13 dimensions, the selected feature is concentrated period of time from 100ms to 150ms and from 250ms to 300ms.

Table 1. Results of classification of single subject

	Subject 1	*Subject 2*	*Subject 3*	*Subject 4*	*Subject 5*
Accuracy	93.2%	87.2%	93.1%	93.8%	86.1%

Table 2. Results of classification of ten subjects (difference methods)

	GMM	*SVM*	*GA+SVM*
Ten subjects	53.8%	82.8%	86.6%

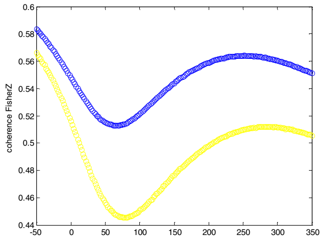

Fig. 3. The event-related coherence (Fisher-Z)(the blue is standard, the yellow is deviation)

3 Conclusions

In this paper we have presented an approach to auditory recognition based on the processing of EEG signals because the event-related coherence of auditory standard and deviation are significant differentiated (see Fig 3). Usually, the EEG signals are recorded from an electrode hat with many electrodes, such as 64, 128, and 256. Even if some electrodes unrelated to the desired task are removed, the number of remaining electrodes may still be large. In order to develop a reliable and efficient EEG signal classification system with less number of electrodes, we designed this experiment, among all fourteen subjects, we choose the relatively stable ten subject data sample and preprocess a total of 947 EEG epochs from two kinds of auditory stimuli. The number of data points in each epoch F4-M2 electrode pair is 302 depends on the event-related coherence. The classification ability of event-related coherence feature set can be measured through classification accuracy. From table 1, the best classification results in these sessions were 93.8% for subject S4, the time taken for the least three trails is lesser than 2s (1800ms). We made significant improvements in the accuracy and speed by employing powerful machine learning algorithms for classification and developing a new dynamic feature in this method. Two-class experiments show that utilization of the event-related coherence parameters as features and genetic algorithm as feature selection improve correct classification at the cost of decreased complexity and computations.

Two bipolar EEG channels prior to the reported multichannel experiment even though in the multichannel experiment no feedback was given. it can be expected that in the latter case the classification accuracy is lower [26] . However, There are shortcomings in my paper, firstly, one difficulty encountered in such a study concerns the lack of published objective comparisons between classifiers [15] [27]. Secondly, one of the major limitations on this research is the lack of the number of the subjects. In the future, find the best parameter configuration and adaptive method for each subject should be investigated. Apart from the current considered base classifiers, the performance of some other classifiers such as k-nearest-neighbor [29] can be further investigated as well. Although our studies were done on healthy subjects, there is a chance that BCI systems such as the one presented in this paper may someday provide potentially the only communication channel for severely disabled people who are otherwise unable to articulate their thoughts and needs[28].

Acknowledgements. Our thanks to supports from the National Natural Science Foundation of China (61171186, 61271345), Key Laboratory Opening Funding of MOE-Microsoft Key Laboratory of Natural Language Processing and Speech (HIT.KLOF.20110xx), and the Fundamental Research Funds for the Central Universities (HIT.NSRIF.2012047). The authors are grateful for the anonymous reviewers who made constructive comments.

References

1. Šťastný, J., Sovka, P., Stančák, A.: EEG Signal Classification: Introduction to the Problem. Radioengineering 12, 51–55 (2003)
2. Curran, E.A., Stokes, M.J.: Learning to control brain activity: a review of the production and control of EEG components for driving brain–computer interface (BCI) systems. Brain Cogn. 51, 326–336 (2003)
3. Jahankhani, P., Kodogiannis, V., Revett, K.: EEG signal classification using wavelet feature extraction and neural networks. In: International Symposium on Modern Computing, pp. 52–57 (2006)
4. Garrett, D., Peterson, D.A., Anderson, C.W., Thaut, M.H.: Comparison of Linear and Nonlinear Methods for EEG Signal Classification. IEEE Transactions on Neural Systems and Rehabilitation Engineering 11, 141–144 (2003)
5. Kaper, M., Meinicke, P., Grossekathoefer, U., Lingner, T., Ritter, H.: Bci competition data set iib: support vector machines for the p300 speller paradigm. IEEE Transactions on Biomedical Engeneering 51(6), 1073–1076 (2004)
6. Pfurtscheller, G., Neuper, C., Flotzinger, D., Pregenzer, M.: Eeg-based discrimination between imagination of right and left hand movement. Electroencephalography and Clinical Neurophysiology 103, 642–651 (1997)
7. Chiappa, S., Bengio, S.: Hmm and iohmm modeling of eeg rhythms for asynchronous bci systems. In: European Symposium on Artificial Neural Networks, ESANN (2004)
8. Millan, J.R., Mourino, J.: Asynchronous BCI and local neural classifiers: An overview of the Adaptive Brain Interface project. IEEE Transactions on Neural Systems and Rehabilitation Engineering, Special Issue on Brain-Computer Interface Technology (2003)
9. Penny, W.D., Roberts, S.J., Curran, E.A., Stokes, M.J.: Eeg-based communication: a pattern recognition approach. IEEE Transactions on Rehabilitation Engeneering 8(2), 214–215 (2000)
10. Pfurtscheller, G., Neuper, C., Schlogl, A., Lugger, K.: Separability of eeg signals recorded during right and left motor imagery using adaptive autoregressive parameters. IEEE Transactions on Rehabilitation Engineering 6(3) (1998)
11. Wang, T., Deng, J., He, B.: Classifying eeg-based motor imagery tasks by means of time-frequency synthesized spatial patterns. Clinical Neurophysiology 115(12), 2744–2753 (2004)

12. Qin, L., Ding, L., He, B.: Motor imagery classification by means of source analysis for brain computer interface applications. Journal of Neural Engineering, 135–141 (2004)
13. Kamousi, B., Liu, Z., He, B.: Classification of motor imagery tasks for brain-computer interface applications by means of two equivalent dipoles analysis. IEEE Transactions on Neural Systems and Rehabilitation Engineering 13, 166–171 (2005)
14. Congedo, M., Lotte, F., Lecuyer, A.: Classification of movement intention by spatially filtered electromagnetic inverse solutions. Physics in Medicine and Biology 51(8), 1971–1989 (2006)
15. Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F., Arnaldi, B.: A Review of Classification Algorithms for EEG-based Brain-Computer Interfaces. Journal of Neural Engineering 4 (2007)
16. Sun, S., Zha, C.: An optimal kernel feature extractor and its application to EEG signal classification. Neurocomputing 69, 1743–1748 (2006)
17. Inouye, T., Shinosaki, K., Iyama, A., Matsumoto, Y.: Localization of activated areas and directional EEG patterns during mental arithmetic. Electroencephalography and Clinical Neurophysiology 86(4), 224–230 (1993)
18. Anderson, C.W., Devulapalli, S.V., Stolz, E.A.: Determining mental state from EEG signals using neural networks. Scientific Programming 4(3), 171–183 (1995)
19. Anderson, C.W., Devulapalli, S.V., Stolz, E.A.: EEG signal classification with different signal representations. In: Girosi, F., Makhoul, J., Manolakos, E., Wilson, E. (eds.) Neural Networks for Signal Processing V, pp. 475–483. IEEE Service Center, Piscataway (1995)
20. Suykens, J.A.K., Van Gestel, T., De Brabanter, J., De Moor, B., Vandewalle, J.: Least Squares Support Vector Machines. World Scientific, Singapore (2002)
21. Neuroscan Edit4.5.American
22. Yang, J., Honavar, V.: Feature subset selection using a genetic algorithm. Intelligent Systems and their Applications, 44–49 (1998); Iowa State Univ.
23. Ren, J., Qiu, Z., Fan, W., Cheng, H., Yu, P.S.: Forward Semi-Supervised Feature Selection. In: Washio, T., Suzuki, E., Ting, K.M., Inokuchi, A. (eds.) PAKDD 2008. LNCS (LNAI), vol. 5012, pp. 970–976. Springer, Heidelberg (2008)
24. Bu, H., Zheng, S., Xia, J.: Genetic algorithm based Semi-feature selection method. In: 2009 International Joint Conference on Bioinformatics, Systems Biology and Intelligent Computing, pp. 521–524 (2009)
25. Oh, I.S., Lee, J.S., Moon, B.R.: Hybrid genetic algorithms for feature selection. IEEE Transactions on Pattern Analysis and Machine Intelligence 26(11), 1424–1437 (2004)
26. Pfurtscheller, G., Neuper, C., Flotzinger, D., Pregenzer, M.: EEG based discrimination between imagination of right and left hand movement. Electroenc. Clin. Neurophys 103(5), 1–10 (1997)
27. Schalk, G., McFarland, D.J., Hinterberger, T., Birbaumer, N., Wolpaw, J.R.: Bci2000: a generalpurpose brain-computer interface (bci) system. IEEE Transactions on Biomedical Engeneering 51(6), 1034–1043 (2004)
28. Thulasidas, M., Guan, C., Wu, J.: Robust Classification of EEG Signal for Brain–Computer Interface. IEEE Transactions on Neural Systems and Rehabilitation Engineering 14(1) (2006)
29. Sun, S., Zhang, C., Zhang, D.: An experimental evaluation of ensemble methods for EEG signal classification. Pattern Recognition Letters 28, 2157–2163 (2007)