Adaboost with SVM-Based Classifier for the Classification of Brain Motor Imagery Tasks

Jue Wang, Lin Gao, Haoshi Zhang, and Jin Xu

Institute of Biomedical Engineering, Key Laboratory of Biomedical Information Engineering of Education Ministry, Xi'an Jiaotong University,
Xi'an 710049, People's Republic of China
juewang@mail.xjtu.edu.cn

Abstract. The Adaboost with SVM-based component classifier is generally considered to break the Boosting principle for the difficulty in training of SVM and have imbalance between the diversity and accuracy over basic SVM classifiers. The Adaboost classifier in the paper trains SVM as base classifier with changing kernel function parameter σ value, which progressively reduces with the changes of weight value of training sample. To testify the validity of the classifier, the classifier is tested on human subjects to classify the left- and right-hand motor imagery tasks. The average classification accuracy reaches 90.2% on test data, which greatly outperforms SVM classifiers without Adaboost and commonly Fisher Linear Discriminant classifier. The results confirm that the proposed combination of Adaboost with SVM classifier may improve accuracy for classification of motor imagery tasks, and have applications to performance improvement of brain-computer interface (BCI) systems.

Keywords: Adaboost; SVM; Classification; Kolmogorov entropy; ERS/ERD; Motor imagery.

1 Introduction

Brain–computer interface (BCI) is a system with direct transmission of information from brain to computer by analyzing the feature of electroencephalogram (EEG), which provides an alternative communication channel independent of brain's normal output pathways of the peripheral nerves and muscles ([1],[2],[3]). EEG recording is simple and noninvasive and can reflect the different brain states and the feature information so that the current BCI technology is mostly based on EEG signals including slow cortical potentials, P300 potentials, visual evoked potential and especially motor imagery EEG recorded from the scalp. Therefore, EEG analysis on discriminating motor imagery tasks in alpha and beta rhythms (i.e., 8-30Hz) over primary sensorimotor cortex is greatly significant for the construction of motor imagery based BCI [4, 5].

Currently, the mostly used methods for classification of left and right hand motor imagination are linear and non-linear classifier such as Mahalanobis distance, Fisher Discriminant Analysis (FDA), neural network etc. Support Vector Machine (SVM)

C. Stephanidis (Ed.): Universal Access in HCI, Part II, HCII 2011, LNCS 6766, pp. 629–634, 2011.

(Vapnik, 1998) is one of the research focuses in classification methods. By using a kernel function to map the training samples from an input space to a high-dimensional feature space, it finds an optimal separating hyperplane in the feature space and controls its model complexity and training error. AdaBoost, one of the most popular technique in machine learning, creates a set of component classifiers by maintaining weights over training samples and adaptively adjusting them after each iteration. In this paper, the Adaboost with RBFSVM as base learner is proposed to apply to discriminate the left and right hand motor imagery tasks. The satisfactory classification results on test data are obtained with the average classification accuracy of 90.2%. Therefore, it has promising potentials to identify the different mental tasks for BCI application.

2 Methodology

2.1 Experiment Data

The hand motor imagination experiment is executed to obtain EEG data with a 1000 Hz sampling rate. Three subjects (male, 24 years old, right-handed) participated in this experiment. EEG data of hand movement imagination was recorded from 32 channels according to the international 10-20. The subjects were seated in front of a screen and asked to imagine either left- or right-hand movement (120 trials, 60 left, 60 right). As shown in Fig.1, Each trial lasted for 14 s. The fixation appeared in the center of the screen as the preparation cue until second 3. Then, a hand, either left or right, appeared on the screen indicating the motor imagery of the corresponding hand for 6 second. The subjects could have a rest from second 9 to second 14. Subjects did not have any feedback during the experiment.

The signal from C3 and C4 over the primary sensorimotor cortex is mainly used in the study because we are interested in the activity of this brain area. More accurately, the signal of C3 and C4 channels after Laplacian filtering are studied in the paper, therefore the neighbors of C3 and C4 are also included.

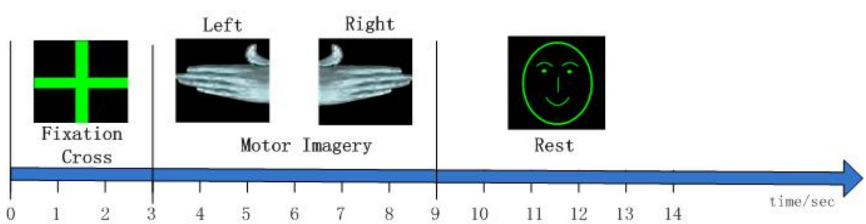

Fig. 1. Timing chart of the experiment

2.2 Feature Extraction by Kolmogorov Entropy

EEG usually oscillates in some particular frequency bands, such as delta (0–4 Hz), theta (4–7 Hz), alpha (8–13 Hz) and beta (14–30 Hz). The motor imagery EEG mainly appears within alpha and beta band. Therefore, the EEG within 8–30 Hz is obtained after band filtering in the paper. Studies by Pfurtscheller showed that the amplitude of alpha and beta rhythms in contralateral hand area would decrease during imagination or preparation for the unilateral hand movement, which was called as ERD (event-related desynchronization); while the amplitude of the corresponding rhythm in ipsilateral hand area would increase, which was called as ERS (event-related synchronization) [6]. The desynchronized EEG process causes the increase of Kolmogorov entropy of EEG over two brain hemispheres, and vice versa. So this paper extracts Kolmogorov entropy as feature vectors characterizing ERD/ERS over left and right hemispheres during motor imagery.

Kolmogorov entropy measures the average rate of information loss and chaotic motion in phase space [7], and quantifies the nonlinear dynamic properties of trajectories in the reconstructed phase space of single channel EEG, which is C3 and C4 in hand motor imagery in the paper. It can be defined as follows: suppose that the d-dimensional phase space is divided into boxed of size ld, the state of the system is now quantified at intervals of time τ [8]. Let Pi0,…,in be the joint probability that $\vec{x}(t=0)$ is in box i0, $\vec{x}(t=\tau)$ is in box i1, …, and etc. According to Shannon, the quantity

$$K_n = -\sum_{i_0 \cdots i_n} P_{i_0 \cdots i_n} \ln P_{i_0 \cdots i_n} \tag{1}$$

is proportional to the information for locating the system on a special trajectory $i_0 \ldots i_n$, if probabilities $P_{i_0 \ldots i_n}$ is known before. Therefore, Kn+1-Kn is the information for predicting in which cell i_{n+1} the system will be. The value of Kn+1-Kn measures the loss of the information of the system from time n to n+1. Thus, Kolmogorov entropy is defined as the average loss rate of information as follows:

$$\begin{aligned} KE &= \lim_{\tau \to 0} \lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{1}{n\tau} \sum_{i=0}^{n-1} (K_{i+1} - K_i) \\ &= -\lim_{\tau \to 0} \lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{1}{n\tau} \sum_{i_0 \cdots i_n} P_{i_0 \cdots i_{n-1}} \ln P_{i_0 \cdots i_{n-1}} \end{aligned} \tag{2}$$

The EEG experimental data within the specific frequency band in 8-30Hz.is divided into 1s segments to extract the Kolmogorov entropy. Finally the relative continuous EEG Kolmogorov entropy time course can be obtained as feature vector.

2.3 Adaboost with RBFSVM-Based Classifiers for Classification

Adaboost is a recent popular developed machine learning method for pattern classification. Freund and Schapire developed the Adaboost model [11], in which the performance of the weak learner can be enhanced effectively by calling the weak or base learning algorithm repeatedly. Each time the training samples with the different distribution are fed into weak learner. The easily classified samples are assigned lower weights and the hard are assigned higher weights, in order to force the base learner to focus on the 'hardest' ones. After many rounds the obtained N weak prediction rules are combined linearly to construct a final classifier, by which the test samples can be classified. The Adaboost algorithm takes a weighted majority vote of all the weak predictions and so the prediction accuracy of the final obtained classifier would be effectively boosted [9, 10].

Developed from the theory of Structural Risk Minimization, SVM is one of the best classifiers since it finds the hyperplane maximizing the separating margin between classes. By using a kernel function to map the training samples from an input space to a high-dimensional feature space, SVM finds an optimal separating hyperplane in the feature space. According to research experiences, SVM classifier with Gaussian Radius Basis Function kernel (RBFSVM) has a satisfactory performance. In order to improve the performance of the classifier, AdaboostSVM, is used to combine RBFSVM classifiers in this paper [11].

The problem is that how to set the parameters, i.e. variance σ for RBFSVM component classifiers during Adaboost iterations. Too large value of σ will result in too weak components classifier, while smaller value of σ will make the strong RBFSVM component overfit the training samples, and result in inefficient boosting because the errors of these component classifiers are highly correlated [11].

In this paper, initially a suitable large value is set to σ, which is the mean standard deviation of all samples. Then RBFSVM with this σ is trained as many cycles as possible as long as more than 50% accuracy can be obtained. The value of σ is decreased slightly by a predefined step. By slightly decreasing the value of σ, we prevent the new RBFSVM component from being too strong for the current weighted training samples, and thus moderately accurate RBFSVM component classifiers are obtained. The reason why moderately accurate RBFSVM component classifiers are favored lies in the fact that these classifiers often have larger diversity than those component classifiers which are very accurate. These larger diversities may lead to a better generalization performance [11]. The detailed procedure is described as follows.

1. Let $Z=(x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n)$ be a set of training samples, where $xi \in X$, $yi \in Y=\{-1,1\}$,represent the feature vectors and classification labels respectively.
2. Initialize the weight of training samples: $w_1(i)=1/n$.
3. Let T be iteration times. For $t=1:T$
   a. Train weak component learner with weighted training samples to obtain hypothesis h_t;
   b. Calculate the training error ε_t

$$\varepsilon_t = \sum_{i=0}^{n} w_t(i) \tag{3}$$

   Where $y(i) \neq h_t(i)$;

   c. Set the weight of weak component learner

$$\alpha_t = \frac{1}{2}\ln\frac{1-\varepsilon_t}{\varepsilon_t} \tag{4}$$

   d. Update the weight of training samples

$$w_{t+1}(i) = \frac{w_t(i)\exp\left[-\alpha_t y_i h_t(x_i)\right]}{D_t} = \frac{w_t(i)}{D_t} * \begin{cases} e^{-\alpha_t}, y_i = h_t(x_i) \\ e^{\alpha_t}, y_i \neq h_t(x_i) \end{cases} \tag{5}$$

   Where D_t is a constant, such that $\sum_{i=1}^{n} w_{t+1}(i) = 1$

e. Output the final hypothesis

$$H(x) = sign[\sum_{t=1}^{T} \alpha_t h_t(x)] \quad (6)$$

3 Results

Based on the feature vectors of Kolmogorov entropy of EEG over two brain hemispheres, the Adaboost classifier with SVM as base learner is applied to discriminate the left and right hand motor imagery tasks for test data. For Adaboost model, the only one parameter i.e the number of iteration T needs to be determined by the predefined small enough training error. Here we choose T=10.

To testify the effectiveness of AdaboostSVM classifier, we select two other classifiers to obtain the classification accuracy for comparison. One is Fisher Linear Discriminant, and the other one is SVM. Fisher linear discriminant is commonly used to classify the different brain state. It is widely recommended to use for the highly accuracy in BCI system. The classification accuracy results of three subjects are showed in Table 1.

For the test data, the results can also be satisfactory with the average classification accuracy of three subjects reaching 90.2%, which highly outperforms Fisher Linear Discriminant (77.6%) and SVM (82.9%).

Table 1. Classification accuracy with different classifier of left- and right-hand movement imagination of three subjects

Subject	Classification Accuracy (%)		
	AdaboostSVM	Fisher	SVM
1	91.7	77.3	83.3
2	85.2	74.1	77.8
3	93.8	81.3	87.5
Average	90.2	77.6	82.9

4 Discussions

In this paper, the Adaboost classifier with RBFSVM as base classifier with changing kernel function parameter σ value, which progressively reduces with the changes of weight value of training sample, is proposed and testified with motor imagery data. Table 1 shows the comparison of classification results on the left and right hand motor imagery tasks by using Fisher Linear Discriminant classifier, and SVM classifier and the Adaboost with SVM as base learner. It can be clearly seen that the results by Adaboost with SVM as base learner is much better than that by SVM only, and SVM is obviously effective than Fisher Linear Discriminant.

Adaboost with RBFSVM-based classifier results in a set of RBFSVM base classifiers with good balance between the distributions of accuracy and diversity over RBFSVM base classifiers and better classification accuracy compared with SVM classifiers. With the ensemble of weighted weak hypotheses, the final classifier is boosted so that the classification performance is effectively boosted. The primary results show that AdaboostSVM classifier can effectively improve the classification performance of motor imagery tasks. Therefore, it could be applied in the mental tasks classification for BCI construction.

Acknowledgments. This work is supported by National Nature Science Foundation (No.: 81071150) and Education Ministry Graveness Science and Technology Cultivation Program (No.: 708080). The authors appreciate the anonymous reviewers for the valuable comments on the paper.

References

1. Pfurtscheller, G., Neuper, C., Guger, C., Harkam, W., Ramoser, R., Schlögl, A., Obermaier, B., Pregenzer, M.: Current trends in Graz brain–computer interface (BCI) research. IEEE Trans. Rehab. Eng. 8, 216–219 (2000)
2. Wolpaw, J.R., Birbaumer, N., Heetderks, W.J., McFarland, D.J., Peckham, P.H., Schalk, G., Donchin, E., Quatrano, L.A., Robinson, C.J., Vaughan, T.M.: Brain–computer interface technology: a review of the first international meeting. IEEE Trans. Rehab. Eng. 8, 164–173 (2000)
3. Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtscheller, G., Vaughan, T.M.: Brain–computer interfaces for communication and control. Clin. Neurophysiol. 113, 767–791 (2002)
4. Pfurtscheller, G., Lopes da Silva, F.H.: Event-related EEG/MEG synchronization and desynchronization: basic principles. Clin. Neurophysiol. 110, 1842–1857 (1999)
5. Pineda, J.A., Allison, B.Z., Vankov, A.: The effects of self-movement, observation, and imagination on mu rhythms and readiness potentials (RP's): toward a brain–computer interface (BCI). IEEE Trans. Rehab. Eng. 8, 219–222 (2000)
6. Pfurtscheller, G., Lopes Da Silva, F.H.: Event-related EEG/MEG synchronization and desynchronization: basic principles. Clin. Neurophysiol. 110, 1843–1857 (1999)
7. Kolmogorov, A.N.: Entropy per unit time as a metric invariant of automorphism. Doklady of Russian Academy of Sciences 124, 754–755 (1959)
8. Heinz, G.S., Wolfram, J.: Deterministic Chaos: An Introduction. WILEY-VCH, Weinheim (2005)
9. Schapire, R.E.: The boosting approach to machine leaming An overview. In: MSRI Workshop on Nonlinear Estimation and Classification, pp. 1–23 (2002)
10. Richard, O.D., Peter, E.H., David, G.S.: Pattem classification, 2nd edn. China Machine Press, China (2003)
11. Li, X.C., Wang, L., Sung, E.: AdaBoost with SVM-based component classifiers. Engineering Applications of Artificial Intelligence 21, 785–795 (2007)
12. Schlogl, A., Neuper, C., Pfurtscheller, G.: Estimating the mutual information of an EEG-based brain-computer interface. Biomed. Tech. 47, 3–8 (2002)