Proceedings of the 2005 IEEE
Engineering in Medicine and Biology 27th Annual Conference
Shanghai, China, September 1-4, 2005

A Statistical Model of Brain Signals with Application to Brain-Computer Interface

Haihong Zhang, Cuntai Guan, Chuanchu Wang
Institute for Infocomm Research, Singapore 119613
{hhzhang,ctguan,ccwang}@i2r.a-star.edu.sg

Abstract

This paper presents a novel approach to improving the robustness of brain-computer interfaces by using a statistical model of brain signals – especially P300. We study the distributions of support vector machine scores for the signals and derive a posteriori probability model of P300/non-P300. We further derive a statistical model for multi-trial brain signals, and apply it to the rejection of undesired signals. Six subjects have been involved in an experimental study. The results demonstrate that the P300 model and the rejection method are appropriate and can help improve the robustness of the system significantly.

1. Introduction

The brain-computer interface (BCI) is an emerging technology for direct communication between brain and computers, which bypasses normal peripheral nerve systems and does not require motor ability. Recent years have seen many BCI systems [1] by detecting event-related potentials in electroencephalogram (EEG).

The P300 potential is a typical event-related EEG signal, elicited in the brain in response to infrequent stimuli such as visual flashes. Farwell and Donchin first demonstrated the use of P300 for BCI in [2]. That system displays a matrix of cells to represent a few letters and commands, and intensifies each row and column successively. When the user's intended cell flashes, a P300 will be generated and can be detected by an appropriate algorithm, and thus the desired letter/command will be recognized.

Many work follows this paradigm. For example, bootstrapping method and data-driven method were investigated for increasing the information transfer rate of P300 speller, respectively in [3] and [4]. Besides, feature extraction methods such as *independent component analysis* were also applied to improving the accuracy for P300 detection [5]. In addition, Guan etc. [6] presented a new paradigm for P300 BCIs, which randomly intensifies each cell instead of a row or a column. They have reported a significantly improved performance on recognition accuracy and information transfer rate.

However, few in the literature have been done to deliberately study the statistics of P300 signals for BCIs. From the BCI viewpoint, we suggest that two statistical issues on P300 are worthy of study. First, how to efficiently capture P300's statistical characteristics despite its high dimensional (spatial-temporal) nature? Second, how to improve BCI systems with the learned P300 statistics?

The aim of this work is to address the above two issues. To tackle the high dimensionality, we first transform P300 signals into their distances (*scores*) to a separate hyperplane (between P300 and non-P300) in a kernel space [7]. A sigmoid function is designed to measure the aposteriori probability of P300 given SVM scores.

We furture derive a probability model of intentional user controls given multi-trial brain signals. Hence by imposing a threshold on the probability, the system is able to reject undesired signals, which may be generated if the user is not well in good control state.

We have conducted experiments on six subjects. The results demonstrate that the P300 model is appropriate and the rejection method can help improve the robustness of the system significantly.

2. Statistical Model of P300/non-P300 Signals

Let's first consider a BCI system in [6]. The system shows a virtual keyboard containing a few buttons representing characters and editing commands. When the BCI starts to work, the buttons are flashed randomly; P300 will be registered in EEG when the target (in focus) button is flashed. There are two categories of signals we need to define: P300 signals (denoted by Θ) and non-P300 signals (denoted by Φ). The Θ category contains all the signals collected when the user focuses on the flashing button. The Φ contains all the other signals.

Fig. (1) displays a graphical representation of P300 and non-P300 signals in terms of mean and standard deviation.

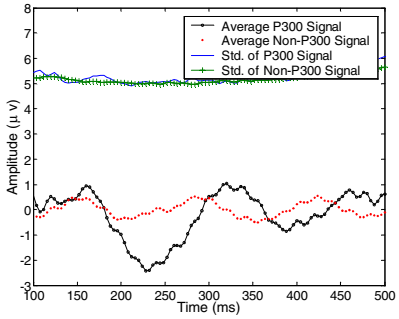


Figure 1. P300 and non-P300 Signals on Channel Fc1. DC component removed.

Apparently the P300 component is present, but the variance is too large across different trials. Hence it would be quite difficult to direct model the statistics of P300/non-P300 signals. Instead, here we resort to studying brain signals in a special space created by support vector machines (SVMs).

SVMs are now a well-known classification method whose principle is to seek maximal margin between two classes. They allow to calculate the distance d from a pattern to the separate hyper-plane, virtually in a high-dimensional nonlinear space,

$$d = h(\mathbf{x}) = \sum_{i=1}^{N} a_i k(\mathbf{x}, \mathbf{x}_i) + b \tag{1}$$

by using a kernel function k (here we uses Gaussian). Here $\mathbf{x}_i$ is one of the N support vectors. In this work, the d space or called *score space* is used as the special space for studying the characteristics of P300/non-P300 signals.

To study the distributions of the SVM scores for real P300/non-P300 signals, we collected data from six subjects and show here the results from a typical subject in Fig. 2. Apparently, on the training set (upper row) the histograms have sharp peaks. The peaks mainly result from the special training strategy for SVMs (see [8]). In [8] Mika also found that on the training set a smaller regularization would produce sharp peaks around the margin, and vice versa. It is also the case in our experiments.

Notably, Fig. (2) suggests that the SVM scores for test patterns are quite close to Gaussian. Therefore, it is reasonable to model the distributions of P300 or non-P300 signals by Gaussian functions. And the distributions can be learned with a separate training set.

Apparently, the variance of P300 and that of non-P300 signals are very similar on each subject. Therefore the two distributions can be described with tied-variance Gaussians. Suppose the learned score distributions are

$$p(d|\Theta) = \mathcal{N}(d - \mu_\theta, \sigma^2), \quad p(d|\Phi) = \mathcal{N}(d - \mu_\phi, \sigma^2) \tag{2}$$

where $\mathcal{N}$ is a normalized Gaussian distribution function, μ denotes the mean and σ^2 the variance. Then we have the posterior probability of signal type Θ given an obtained SVM score d:

$$\begin{aligned} P(\Theta|d) &= \frac{p(d|\Theta)P(\Theta)}{p(d|\Theta)P(\Theta) + p(d|\Phi)P(\Phi)} \\ &= \frac{1}{1 + rp(d|\Phi)/p(d|\Theta)} \\ &= \frac{1}{1 + r\exp(ad + b)} \end{aligned} \tag{3}$$

where r is the ratio $P(\Phi)/P(\Theta)$, and a and b are given by

$$a = \frac{\mu_\theta - \mu_\phi}{\sigma^2} \qquad b = \frac{\mu_\phi^2 - \mu_\theta^2}{2\sigma^2} \tag{4}$$

The other posterior probability is easy to calculate therein:

$$P(\Phi|f) = 1 - P(\Theta|f) \tag{5}$$

3. Statistical Modeling of Multi-trial Brain Signals

P300 BCI systems usually perform a recognition after receiving a few trials where each trial may contain a round of intensifications on each button. Here we will use the above sigmoid model to distinguish two types of the multi-trial signals: control brain signals Ξ (called control patterns) and non-control brain signals Ψ (called non-control patterns). Control patterns are generated as the user is consciously controlling the interface, while non-control patterns are generated when the user is at idle state or performing other action with no intention to control the interface.

In a multi-trial case, a brain signal pattern for N_r trials and N_s targets (e.g. control commands) can be illustrated as in Fig. 3(a). Each S_i^j represents the brain signal in response to a particular stimulus. When the user is really sending control signals, he/she is supposed to keep focus on the corresponding button on the screen for a certain period. The brain signals thus may show the pattern Ξ with a single row of Θ signals (see Fig. (3(b))).

On the contrary, a Ψ pattern does not have a single row of Θ type signals. For example, such a pattern would be present when the user is skimming across the screen so Θ patterns would be randomly distributed in different rows.

Let's consider the probability of control pattern Ξ given collected SVM scores D for the multi-trial. The S would contain a single row of Θ patterns, but the particular row is unknown. We use $P(R_i)$ to denotes the probability that the Θ patterns lie in the R_i row. Then we have

$$P(\Xi|D) = \sum_{i=1}^{N_s} P(\Xi, R_i|D)P(R_i) \tag{6}$$

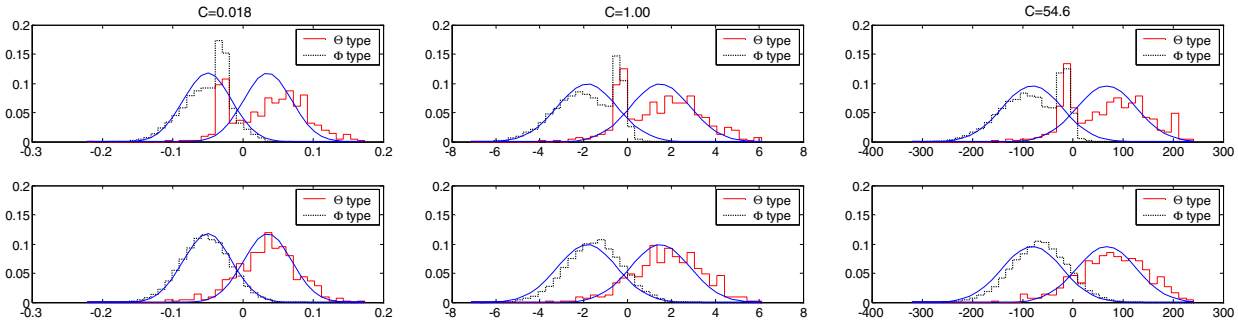
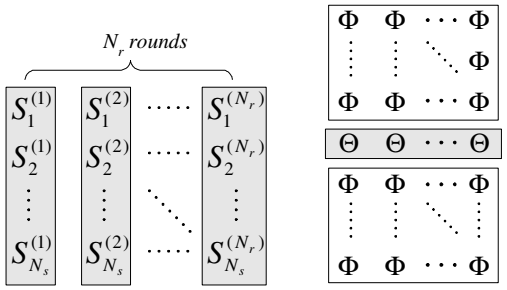


Figure 2. P300 SVM Score Distributions over different regularization factor. We show here the histograms of the scores d on training set (upper row) and on test set (lower row) respectively, and also superpose the Gaussian approximations for training sets in both rows.



(a) Brain Signals (b) A Control Pattern

Figure 3. Brain Patterns (Ξ)

where N_s represents the number of total rows (i.e. the number of buttons). For simplicity, one can set $P(R_i) = \frac{1}{N_s}$ in practice, which means that each control command is equally likely to be sent by the user.

It is straightforward to derive

$$P(\Xi, R_i|D) = P(d_{ij} \in \Theta, d_{kj} \in \Phi, \forall i, j \in [1 \cdots N_r], \forall k \neq i|D) \quad (7)$$

where d_{ij} denotes the SVM score of $S_i^{(j)}$ in Fig. (**??**). Assume the signals $S_i^{(j)}$ are independently generated by the brain. It follows that

$$P(\Xi, R_i|D) = \prod_j P(\Theta|d_{ij}) \prod_{k,j,k\neq i} P(\Phi|d_{ij}) \quad (8)$$

So far we have proposed a mathematical method (using Eq. (6) and Eq. (8)) to calculate the probability $P(\Xi|D)$. Using a threshold on the probability would allow the system to reject non-control patterns from being recognized. In fact, the threshold would also filter and discard some "bad" patterns of brain signals that would be present when e.g. the user is not in good control state.

Hence it is critical to set a threshold to reject ideally all the undesired brain signals, leaving true control patterns to be processed by the recognition machine. In this work we use an empirical method to set the threshold.

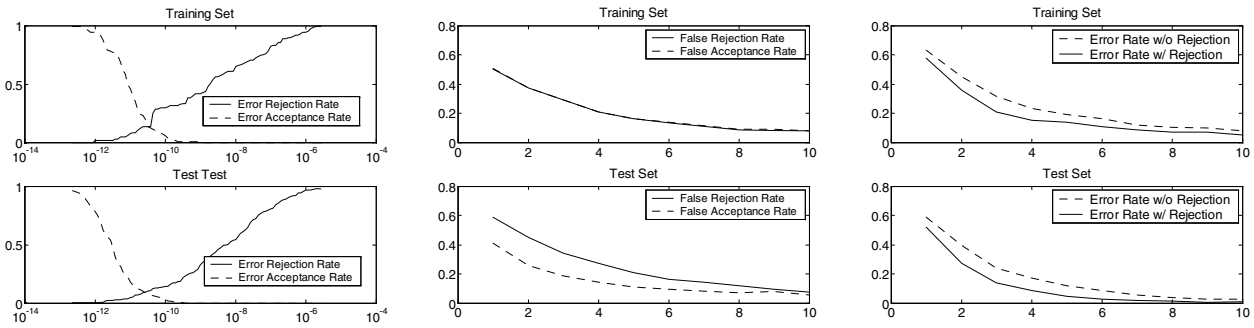
4. Experimental results

The data were collected on six subjects for offline analysis. In particular, we would like to study a special type of non-control patterns called *varying-focus patterns*, which are generated when the user changes focus from trial to trial. The varying-focus patterns are more close to true control patterns than others, since they also contain P300 signals in each trial. Therefore, it is more challenging to reject these patterns than the others. Besides, they are also easy to simulate(by permuting true multi-trial signals so that P300 signals are distributed into different rows in the matrix in Fig. (3(a))).

Our experiments with the true control signals and the simulated varying-focus signals are illustrated in Fig. (4). In particular, Fig. (4(a)) draws the rejection accuracy as a function of threshold (on a typical subject, rejection based on five trials). The results are as expected: larger thresholds tend to produce less false acceptance (FA) or more false rejection (FR). Hereafter we always choose the threshold that achieves *equal error rate* on the FA/FR curves.

Fig. (4(b))shows averaged false rejection errors on the six subjects. It is clear that with more trials (i.e. more information) the system would have higher rejection accuracy.

Another interesting study is on the the influence of rejection on recognition. Specifically, we applied the recognition machine (here we use SVMs; see [6]) onto the remaining brain signals after rejection. The results are illustrated in Fig. (4(c)). Table 1 also lists some detailed error rates (with 10 trials). It is evident that the application of rejec-



(a) Rejection ROC (b) Rejection vs # Trial (c) Recognition vs Rejection

Figure 4. Rejection Experiment Results

	w/o Rej.	w/ Rej.	Improvement
Training Set	7.9%	5.1%	35.4%
Test Set	2.4%	0.9%	62.5%

Table 1. False Recognition vs Rejection

tion can considerably improve the recognition correct rate. In other words, when the rejection mechanism is applied, most incorrectly-recognized patterns will be rejected at the cost of only a few false rejections on correctly-recognized patterns.

We have also conducted a few online tests and obtained initial results that demonstrate the system's robustness against undesired brain signals in online tasks. Those results are yet to be analysed and dicussed.

5. Conclusion

In this work we have studied the statistics of P300 signals, and applied it to improving the robustness of a BCI system by rejecting undesired signals. Based on an aposteriori probability model of P300, we developed a statistical model for discribing the likelihood of multi-trial brain signals with respect to intentional user controls, and designed a rejection mechanism on the likelihood. Our extensive experimental study suggests that the presented P300 model is appropriate; it also suggests that the rejection mechanism is useful for improving BCIs that are based on the detection of event-related potentials . The undesired signals would confuse, or cause false actions in, many existing BCI systems. Thus rejecting those signals would lead to new systems with considerably improved robustness.

Acknowledgement

The authors are thankful to Prof. Klaus-Robert Müller for his valuable comments on this work. The authors also thank Dr. Li Yuanqin for helpful discussions and Dr Manoj Thulasidas for helping experimental work.

References

[1] J.R. Wolpaw, N. Birbaumer, D.J. MacFarland, G. Pfurtscheller, and T.M. Vaughan. Brain-computer interface for communication and control. *Clinical Neurophyiology*, 113:767–791, 2002.

[2] L. A. Farwell and E. Donchin. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potential. *Electroencephalography and Clinical Neurophysiology*, 70:510–523, 1988.

[3] E. Donchin, K. M. Spencer, and R. Wijesinghe. The mental prosthesis: assessing the speed of a p300-based brain-computer interface. *IEEE Transactions on Rehabilitation Engineering*, 8:174–179, 2000.

[4] P. Meinicke, M. Kaper, F. Hoppe, M. Heumann, and H. Rotter. Improving transfer rates in brain computer interfacing: A case study. In *Advances in Neural Information Processing Systems*, 2002.

[5] N. Xu, X. Gao, B. Hong, X. Miao, S. Gao, and F. Yang. Bci competition 2003ł data set iib: Enhancing p300 wave detection using ica-based subspace projections for bci applications. *IEEE Transactions On Biomedical Engineering*, 51(6):1067–1072, 2004.

[6] C. Guan, M. Thulasidas, and J. Wu. High performance p300 speller for brain-computer interface. In *IEEE Int. Workshop on Biomedical Circuits and Systems (BioCAS)*, 2004.

[7] K.R. Muller, S. Mika, G. Rastch, K. Tsuda, and B. Scholkopf. An introduction to kernel-based learning algorithms. *IEEE Transactions on Neural Networks*, 12(2):181–202, 2001.

[8] S. Mika. *Kernel Fisher Discriminants*. PhD thesis, University of Technology, Berlin, 2002.