ROBUST CLASSIFICATION OF EVENT-RELATED POTENTIAL FOR BRAIN-COMPUTER INTERFACE

Cuntai Guan[1], Xiaoyuan Zhu[1], Manoj Thulasidas[1], Sitaram Ranganatha[1], Jiankang Wu[1]

[1] Institute for Infocomm Research, 21 Heng Mui Keng Terrace, Singapore 119613
{ctguan,stuxz2,ranga,manoj,jiankang}@i2r.a-star.edu.sg,

Abstract: Brain-computer interface (BCI) adds a new dimension to human-computer interface, especially for patients suffering from complete paralysis or locked-in syndrome. In this paper, we present our research results on a P300 based BCI system. We are particularly interested in the situation where only small number of trials or even a single trial is available. Machine learning approach is adopted in our system. In order to enhance the classification performance, we propose a new feature extraction method. Experimental results show that the new features have improved classification accuracy, and increased the information transfer rate as well.

Keywords: P300 evoked potential, brain-computer interface, classification, information transfer rate.

INTRODUCTION

Brain-computer interface (BCI) provides a direct connection between human brain and computer. It does not rely on the muscles or brain's normal output pathways of peripheral nerves [1]. It adds a new dimension to existing human-computer interface, especially for people who do not have access to normal channels of man-machine communication and control. BCI research has attracted significant attentions in the past years. For detailed information regarding BCI, readers are referred to literature [1-6]. Various kinds of signals can be extracted from EEG to carry out BCI tasks, for example, slow cortical potential [4]; m and b rhythms [5]; EEG (de)synchronization evoked by motor imagery [6]; P300 evoked potential [7], etc. This paper presents research based on the paradigm proposed by Farwell and Donchin in [7], called P300 speller.

When a subject is presented with a visual, audio, tactile or other sensory stimulus, certain evoked signal can be detected in EEG, referred to as event-related potential (ERP). P300 potential is one of the well studied and most stable potentials. It is elicited by an oddball paradigm. That is, an anticipatory event creates a measurable potential difference at the central sites of EEG measurements. This positive potential typically occurs around 300 milliseconds after the event occurs.

The P300 speller proposed in [7] is based on visual stimulus. The user is presented with a matrix of characters or words. The user's task is to focus his attention on characters or words of interest in the matrix. Rows and columns of this matrix were successively and randomly intensified for some time (usually a few hundred milliseconds). The potential around 300ms evoked by those stimuli corresponding to rows and columns containing desired characters or words are stronger than those evoked by stimuli that do not contain the desired characters or words. The usual practice is to repeat the intensification for several times (typically 15 to 30) in order to reliably detect the correct row or column which contains the intended character or word. The key challenge here is to reduce the number of repetitions while still detecting P300 potential reliably.

The traditional method of detecting P300 potential is to average multiple repetitive occurrences of EEG data for a specific stimulus. Typical classification methods based on the averaging data are peak picking, stepwise discriminant analysis, area and covariance [7]. These averaging methods implicitly assume that the amplitude and latency of individual P300 ERPs are constant from trial to trial. Unfortunately this is not true [8-11]. To tackle these issues, it is required to estimate P300 potentials for single trials. In [6], the authors propose an ML method to estimate P300 potential by taking into consideration the possible trial-to-trial latency changes, but assume the amplitudes do not change. [7] further improves the estimation by relaxing the constant amplitude constraint in [6]. [8] proposes a subspace regularization method to estimate single trial potential for single channel and the authors later expand the method to multiple channel case [9]. None of these single trial estimation methods are tested on classification tasks for BCI. In [12], machine learning approach is proposed in P300 speller classification. It achieves both good classification accuracy and high information transfer rate. This approach outperforms all EEG-based BCI-system thus far [12]. Our P300 speller is built upon support vector machine (SVM) classification framework.

One of the central challenges in developing brain-computer interface is to increase the information transfer rate. Information transfer rate is represented as:

$$B = \frac{60}{T}(\log_2 N + P\log_2 P + (1-P)\log_2(\frac{1-P}{N-1})) \quad (1)$$

where rate B is in bits per minute; T is the average inter-command interval in seconds – proportional to the number of repetitions of row/column intensification; N is the number of choices in the classification task (number of characters or words in P300 speller); and P is the probability for classification. From (1) it can be seen that, in order to increase *B*, one can either increase *P* (maximum 100%) or decrease *T*, or do both.

Our goal is to boost the classification accuracy *P* for smaller *T*. This was achieved by incorporating several techniques. First, a baseline system was built with SVM and artefact removal. Then, we conducted a thorough study on the relationship between classification accuracy and the frequency components of P300 signal, so as to find the effective frequency band for P300 signal directly. Finally we proposed a new feature for the enhancement to the existing feature vectors.

The paper is organized as follows: In next section, a brief description of P300 speller and related signal processing and classification methods is first given. Dataset for the experiments and baseline classification results are presented in the second half of this section. In following section, we discuss the model of EEG and P300 potential from the classification point of view. This is the first attempt to reveal the relationship between spectrum components of P300 signals and the classification accuracy. In the next section, we describe the definition of the new feature, and feature selection method, and give experiment results.

P300 SPELLER

P300 Speller Paradigm

In the P300 speller, a subject is presented with a six by six matrix of characters (see Figure 1). The subject's task is to focus his attention on alphanumeric symbols in the matrix, one at a time. Rows and columns of this matrix were successively and randomly intensified for 100 milliseconds (ms), followed by 75 ms of non-intensification. That is, one complete round of display takes 2.1 seconds. Two out of twelve intensifications of rows or columns contain the desired character (i.e., one particular row and one particular column). The responses evoked by these two infrequent stimuli are different from those evoked by the stimuli that did not contain the desired character [7]. Specifically, the positive potentials for infrequent (targeted) stimuli are relatively higher than that for frequent stimuli. Our task is to identify which stimulus evokes the most significant positive potential around 300ms after stimulus onset so as to recognize the target symbols.

Signal Processing and Classification

The EEG data are sampled at 250Hz. We take the segment between 250-500ms after the onset of the intensification (stimulus) as an epoch. For each epoch, the data are first fed into a low-pass filter and down-sampled with 60Hz sampling rate using a moving-averaging window. A ten-order Chebyshev II type IIR filter is used. The passband cut-off frequency of this filter is 10Hz. We manually select 25 channels around Cz, CPz and FCz positions from all 64 channels. A feature vector is formed by concatenating data from selected channels and sent to a classifier for classification. The feature vector is constructed as follows:

$$x = [x(1)^T, \cdots, x(n)^T, \cdots, x(N)^T]^T \qquad (2)$$

where *T* denotes matrix transpose, and *N* is the time index, and *x(n)* is a *K*-channel vector at time instance *n*, $x \in R^{N\times K}$.

We use support vector machine (SVM) [13] for classification. SVM is a powerful approach for pattern recognition. It does not need any distributional assumptions about the data while providing very good discriminative solution and generalization at the same time. It is pretty suitable for our case, where we normally only have a relatively small amount of training data due to the non-trivial difficulties in data collection efforts.

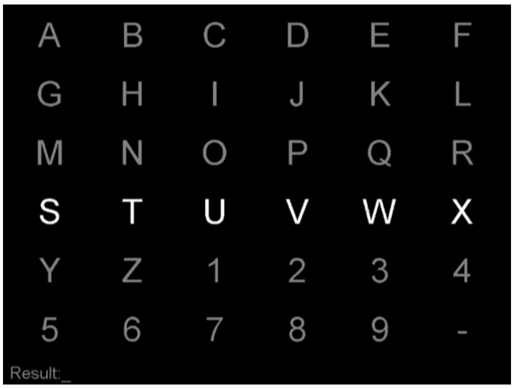

Figure 1. P300 Speller Display. Rows and columns are highlighted randomly and are repeated several rounds. The picture shows the 4th row is intensified.

In the case of single trial classification, the decision function for an SVM classifier is in the form of:

$$f(x) = \mathrm{sgn}(\sum_{i=1}^{m} y_i a_i k(x, x_i) + b) \qquad (3)$$

where $x \in R^{N\times K}$ is the test data, $x_i \in R^{N\times K}$ are the training data, $y_i \in \{-1,1\}$ are the class labels and $k()$ is the kernel function. We use Gaussian kernel function, which is a reasonable choice for EEG signal, as filtered EEG waveforms are used as input feature of the classifier and the EEG waveforms are best modelled with Gaussian distributions. Instead of using (3) to directly classify each epoch into positive or negative class, we sum up the SVM margins for all repetitions of the same stimulus and then make a decision for rows and columns respectively. For example, after a number of rounds of stimuli are presented, we decide the most probable column and row by the following decision functions:

$$n_r = \underset{n=1,\cdots,N_r}{\arg\max}\left\{\sum_{j=1}^{R}\left[\sum_{i=1}^{m} y_i a_i k(x_{r,j}^{(n)}, x_i) + b\right]\right\}$$
$$n_c = \underset{n=1,\cdots,N_c}{\arg\max}\left\{\sum_{j=1}^{R}\left[\sum_{i=1}^{m} y_i a_i k(x_{c,j}^{(n)}, x_i) + b\right]\right\} \quad (4)$$

where N_c and N_r are the number of columns and rows in the alphanumeric matrix (both are 6 in our case), *R* the number of repetitions, $x_{c,i}$ and $x_{r,i}$ are the test data for columns and rows respectively. Once a column index and a row index are determined, a character is recognized. Here we do not average the signal in the data space but in the classification score space. This ensures that once a classifier is trained, it can be used online for a P300 speller with an arbitrary number of rounds. Usually, SVM requires a balanced number of training data among various training classes. However, in our case, we do not need to make a decision for every individual row or column. We only need to make a decision among columns or rows, Therefore, we can make full use of our training data. This can potentially improve the SVM model, because among training data, there are 6 times more elements without targeted stimuli than those with targeted stimuli.

Database

We collected data from two subjects. The data sets are labelled as A and B respectively. Each set contains 31 occurrences of characters. The number of repetitions of intensification for each character is 10. Since for each round of intensification, one row and one column contains the intended character, so we have 620 trials of P300 data, and 3100 trials of non-P300 data for each of the two datasets. During our evaluation hereafter, we use leave-one-out approach to test each character. To evaluate the system better, we add another dataset from BCI 2003 Competition [14], denoted as dataset C, which was collected with the similar paradigm as the one depicted in figure 1. The only difference for C is that it is sampled with a slightly different sampling rate (240Hz). In dataset C, there is a training set as well as a test set. The training set contains 42 occurrences of characters and the test set contains 31 occurrences of characters. For each character, the number of intensifications is 15. For dataset C, we train models using the training set and evaluate it with the test set.

Baseline Results

The accuracy for various numbers of repetitions is calculated in this way, when the number of repetition is *L*, we use every successive *L* trials to make a classification decision based on function (3). For instance, if a character is repeated by 15 rounds of intensification, we get 15 single-trial tests for this character, 14 two-trial tests, and so forth.

The character classification results are listed in table 1 for three datasets. As we are interested in small number of trails, we only list results for repetitions 1 to 6. When the number of repetitions increases, the accuracy will increase monotonically. Typically, for a number of repetitions above 10, we get 100% accuracy. It is worthwhile to mention that, for dataset C, the data is manually screened to remove any EOG artefacts (noise caused by eye blinking), so it is “clean”. But in dataset A and B, about 5% of trials contain EOG artefacts. Instead of rejecting them from the evaluation set, we use an automatic EOG removal algorithm to reduce its adverse effect [15]. In fact, dataset C has better accuracy than A and B. EOG could be one of the reasons.

%	Number of Character Repetitions						Avg
Subj	1	2	3	4	5	6	
A	34.6	42.3	48.3	51.3	63.1	73.1	52.1%
B	44.4	59.6	60.9	67.2	73.3	77.5	63.8%
C	40.4	57.9	75.2	83.8	88.6	93.5	73.2%

Table 1: Baseline character accuracy for three subjects with 1 to 6 repetitions.

FREQUNCY COMPONENTS OF P300 POTENTIAL SIGNAL

EEG signal is generated due to synchronous synaptic firing of the brain cells. The major frequency components of EEG are in the range of 0.1-30Hz, whereas event evoked potential is considered as a slow change waveform composed of low frequency components. EEG signal can be represented with the following model:

$$x(n) = s(n) + e(n) \quad (5)$$

where $s(n)$ is the P300 potential, $e(n)$ is the background EEG. It is well recognized that $s(n)$ is composed of low frequency components. However, since the spectrum of $s(n)$ and $e(n)$ are overlapped, it is impossible to directly separate of $s(n)$ from $e(n)$ in practise. Therefore it is difficult to explicitly observe the actual frequency components of $s(n)$. Here we try to reveal the P300 frequency components with the aid of classifier, so that we can look at the problem from a different angle. We evaluate the classification accuracy for all three subjects by setting the cut-off frequency (where -6dB attenuation occurs) of the low-pass filter from 1Hz to 27Hz. The classification results for one of the subjects are depicted in figure 2. The results for other subjects show very similar pattern.

In figure 2, the numbers besides the legend denote the number of repetitions for that test. It can be observed that when the cut-off frequency of the low-pass filter changes between 8 to 27Hz, the classification accuracy maintain at the similar level. But when the frequency is below 8Hz, the accuracy drops dramatically. We can conclude from these observations that, P300 potentials mainly span 0-8 Hz in spectrum. Another observation from the results is that the classification accuracy (especially for repetition 2 or more) shows a slightly decreasing trend when the cut-off frequency increases. The reason for this trend is that, as the cut-off frequency increases, more background EEG data are included and they are obviously noise for our task.

All the experiments in this paper, unless otherwise stated, are done with cut-off frequency of 10Hz.

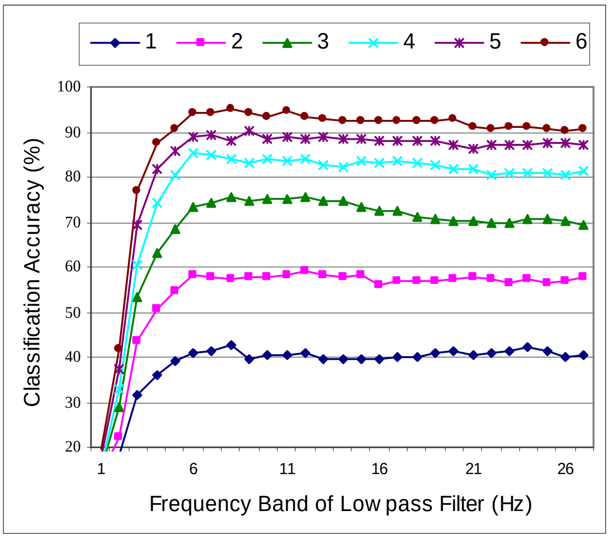

Figure 2. Classification accuracy of subject C with various cut-off frequency of low-pass filtering. The numbers besides the legends denote the number of repetitions used for each test.

DYNAMIC FEATURE TO ENHANCE CLASSIFICATION ACCURACY

Dynamic Feature Calculation

The amplitude of P300 potential for each trial is relatively low compared to the on-going EEG. There are various sources of noises added to EEG signal. This makes the differentiation of P300 trials and non-P300 trials rather difficult. The feature normally used for P300 potential detection reflects the static shape of the P300 temporal waveform. To enhance the distinguishability between P300 and non-P300 signal, here we propose to make use of the temporal transition information. Temporal derivative feature was successfully deployed in speech recognition [16], where the first and second order derivatives are used to improve robustness of speech recognition.

Given the static temporal feature in (2), the time derivative of feature $x(t)$ at time *t* is:

$$\Delta x(t) = \frac{\partial}{\partial t} x(t) \tag{6}$$

Since $x(t)$ is a discrete time representation, simply using a first order or second order difference is inappropriate to approximate the derivative because it is very noisy [17]. It is therefore approximated by an orthogonal polynomial fit as follows [16]:

$$\Delta x(n) \approx a \sum_{m=-M}^{M} m x(n+m) \tag{7}$$

where a is a normalization constant, and the computation is performed over a window of $2M+1$. This is actually a least square estimate.

After calculating the dynamic feature, the final feature vector sent to SVM classifier is an augmented one:

$$x = [x(1)^T, \cdots, x(N)^T, \Delta x(1)^T, \cdots, \Delta x(N)^T]^T \tag{8}$$

Interpretation of Dynamic Feature

The dynamic feature can be expressed as:

$$\Delta x(n) \approx a \sum_{m=-M}^{M} m x(n+m) = w(n) \otimes x(n) \tag{9}$$

where $\otimes$ denotes convolution and $w(n)$ is the weighting function which is defined as follows:

$$w(n) = \begin{cases} a\,n & n = -M, \cdots, 0, \cdots, M \\ 0 & \text{others} \end{cases} \tag{10}$$

$\Delta x(n)$ is interpreted as the output of an FIR filter defined by $w(n)$.

EXPERIMENTS

Classification Performance Improvement by Using Dynamic Feature

The classification results after incorporating the new dynamic feature are given in table 2. Compared to the baseline results in table 1, the average accuracy improvements for three subjects are 6%, 6.4% and 3.2% respectively.

Figure 3 shows detailed accuracy improvement for various cut-off frequency of the low-pass filtering. It can be seen that the improvement is achieved for all frequency. Therefore this new feature makes the classification more robust against the use of low-pass filters.

%	Number of Character Repetitions						Avg
Subj	1	2	3	4	5	6	
A	39.7	46.6	54.4	62.8	68.5	76.9	58.1%
B	45.9	64.2	71.4	76.6	80.0	83.3	70.2%
C	45.5	63.2	79.6	85.6	90.1	94.1	76.4%

Table 2: Classification accuracy using dynamic feature for three subjects with 1 to 6 repetitions for each character.

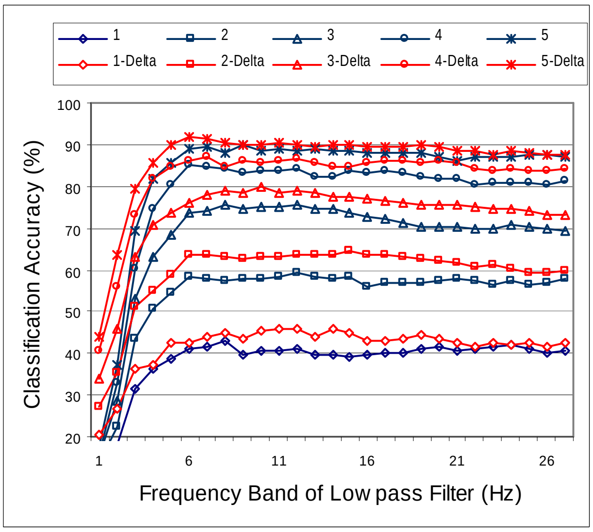

Figure 3. Accuracy improvement for various cut-off frequencies of the low-pass filter for subject C.

Effect of FIR Filter Order for Dynamic Feature Calculation

There are two parameters in calculating dynamic features, order of the FIR filter 2M+1 and the normalization constant a . Since the feature vectors are normalized for each individual dimension before they are sent to an SVM for classification, the effect of a is actually eliminated. We only need to study the effect of the FIR filter order. By varying the order from 3 to 9 (corresponding time windows of 50 – 150ms), the classification results for various subjects are listed in Table 3.

From these results we can see that for subject C, the accuracy for different order of FIR filter is almost the same, while for subject A and B, the accuracy changes for different order, and the best order is 5. The individual accuracy for subject C is also shown in Figure 4

%	Order of FIR Filter			
Subj	3	5	7	9
A	57.85%	58.15%	58.03%	57.90%
B	70.20%	70.23%	70.11%	69.78%
C	75.69%	76.41%	76.46%	76.39%

Table 3: Classification accuracy using dynamic feature for three subjects with 1 to 6 repetitions for each character. The accuracy in the table is the average over six repetitions.

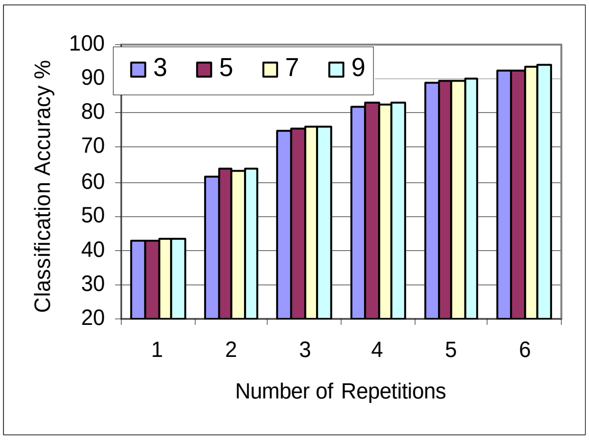

Figure 4. Accuracy for various filter order. The numbers besides the legend symbols denote the FIR filter order. This result is for subject C.

CONCLUSION

In this paper, we present the methods used in our BCI system, P300 speller. This system adopts a single-trial based classification algorithm built on powerful machine learning method, support vector machine, which gives very good classification and generalization performance. The classifier is trained with individual trials of EEG data instead of averaging data. This way, our system is rather robust to latency variability, one of the big problems in P300 speller when averaging method is deployed. The reason is that latency invariability is the basis of averaging approaches, while for single-trial approaches, the classifier can learn the latency variability from the data.

We propose a new dynamic feature to enhance the P300 speller performance. The new feature is shown to improve the classification accuracy. The average relative improvements for the three subjects in our experiments range from 4.4% to 11.5% when first order dynamic feature is used. In our future research, we will study if the second or higher order temporal derivative can be used to improve the accuracy further.

Our study also reveals the frequency components of P300 signal from a different angle. A low pass filter of cut-off frequency around 10Hz is recommended for optimal performance.

We are now carrying out on-line experiments to test the accuracy and robustness of our system with more subjects. In these experiments, we asked subjects to purposely generate various artefacts (such as frequent eye blinking, body movement, talking, drinking, singing, laughing, etc) in order to see the robustness of the system. Preliminary result showed all those artefacts actually have very little effect on the classification accuracy. Detailed results on these experiments will be reported elsewhere.

ACKNOWLEDGEMENT

This work is funded and supported by the Agency for Science and Technology (A*Star), Singapore.

We thank Wadsworth Center, Albany, for sharing BCI dataset (P300 evoked potentials) acquired using BCI2000 P3 Speller Paradigm.

REFERENCES

[1] Wolpaw J R, Birbaumer N, MacFarland D J, Pfurtscheller G and Vaughan T M (2002): 'Brain-computer interface for communication and control'. *Clin. Neurophyiol*., 113:767-791, 2002.

[2] Kubler A, Kotchoubey B, Kaiser J, Wolpaw J, Birbaumer N (2001): 'Brain-computer communication: Unlocking the locked in', *Psychol. Bull.*, 127(3):358-375, 2001.

[3] Curran E A, Stokes M J (2003): '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.

[4] Birbaumer N, Hinterberger T, Kübler A and Neumann N (2003): 'The Thought-Translation Device (TTD): Neurobehavioral Mechanisms and Clinical Outcome**'**, *IEEE Trans on Neural Systems and Rehab Eng*, 11(2):120-123, June 2003.

[5] McFarland D, Sarnacki W, Wolpaw J(2003): 'Brain-computer interface (BCI) operation: optimizing information transfer rates', *Biol. Psychol.*, 63:237-251, 2003.

[6] Muller-Gerking J, Pfurtscheller G, Henrik Flyvbjerg (1999): 'Designing optimal spatial filters for single-trial EEG classification in a movement task'. *Clin. Neurophysiol.*,110:787-798, 1999.

[7] Farwell L A and Donchin E (1988): 'Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potential', *Electroencephalogr and Clin Neurophy*, 1988, 70:510-523.

[8] Pham D T, M¨ocks J, K¨ohler W, and Gasser T (1987):'Variable latencies of noisy signals: Estimation and testing in brain potential data', *Biometrika*, vol. 74, pp. 525–533, 1987.

[9] Ja´Skowski P and Verleger R (1999):'Amplitudes and Latencies of Single-Trial ERP's Estimated by a Maximum-Likelihood Method', *IEEE Trans on Biomedical Eng*, 46(8):987-992, August 1999.

[10] Karjalainen P A, Kaipio J P, Koistinen A S, Vauhkonen M (1999):'Subspace Regularization Method For The Single-Trial Estimation Of Evoked Potentials', *IEEE Trans on Biomedical Eng*, 46(7):849-860, July 1999.

[11] Ranta-Aho P O, Koistinen A S, Ollikainen J O, Kaipio J P, Partanen J and Karjalainen P A (2003): 'Single-Trial Estimation of Multichannel Evoked-Potential Measurements', *IEEE Trans on Biomedical Eng*, 50(2):189-196, February 2003.

[12] Meinicke P, Kaper M, Hoppe F, Heumann M and Rotter H (2002):'Improving Transfer Rates in Brain Computer Interfacing: A Case Study', *in Proc NIPS* 2002, Canada.

[13] Vapnik V N (1998):'*Statistical Learning Theory'*, John Wiley and Sons, New York. 1998.

[14] BCI Competition 2003 Web Site: http://ida.first.fraunhofer.de/projects/bci/competition/

[15] Ranganatha S, Guan C, Thulasidas M, Xu W, Zhu X, Wu J (2004): 'Comparison of Artifact Removal Methods on Their Effect on Motor Task and Imagery Classification in a Brain-computer Interface', *in MEDSIP2004, Malta*.

[16] Furui S (1986):'Speaker-independent isolated word recognition using dynamic features of speech spectrum**',** *IEEE Trans on ASSP*, 34 (1):52-59, Feb 1986.

[17] Rabiner L and Juang B-H (1993): 'Fundamentals of speech recognition', Prentice Hall, 1993.