Robust Classification of EEG Signal for Brain–Computer Interface

Manoj Thulasidas, *Member, IEEE*, Cuntai Guan, *Senior Member, IEEE*, and Jiankang Wu

Abstract—We report the implementation of a text input application (speller) based on the P300 event related potential. We obtain high accuracies by using an SVM classifier and a novel feature. These techniques enable us to maintain fast performance without sacrificing the accuracy, thus making the speller usable in an online mode. In order to further improve the usability, we perform various studies on the data with a view to minimizing the training time required. We present data collected from nine healthy subjects, along with the high accuracies (of the order of 95% or more) measured online. We show that the training time can be further reduced by a factor of two from its current value of about 20 min. High accuracy, fast learning, and online performance make this P300 speller a potential communication tool for severely disabled individuals, who have lost all other means of communication and are otherwise cut off from the world, provided their disability does not interfere with the performance of the speller.

Index Terms—P300, brain–computer interface, event related potential, speller, support vector machine (SVM).

I. INTRODUCTION

BRAIN–COMPUTER interface (BCI) [1], [2] provides a direct communication channel from the user's brain to the external world by reading the electrical signatures of brain's activity and its responses to external stimuli. These responses can then be translated to computer commands, which can either be carried out or made known to others, thus providing a communication link, particularly for people with severe disabilities. The input to our BCI system described here is the electroencephalogram (EEG) signals. EEG signals, however, suffer from high noise levels due to the low conductivity of the human skull. In order to identify patterns from EEG signals for the speller application, we use a support vector machine (SVM) [3], [4] as our classifier. Although SVM classifiers have been used [5]–[7] for the same purpose before, high online accuracies have not been reported.

One criterion for judging the usefulness of a BCI system is whether it can be used in an online scenario. Usability of a BCI system further depends on several factors, such as its accuracy, how fast it can be trained, how long it takes to prepare the subject, and so on. These issues are especially critical, since a typical subject is physically disadvantaged. This article addresses these key issues.

Fig. 1. User display in P300 Speller.

II. METHODOLOGY

A. P300 Speller

The speller system studied in this article is based on the P300 event-related potential that is elicited by an oddball paradigm. The P300 potential is created in the central sites of EEG measurements when an infrequent and anticipated event occurs. P300 is the signature of the user's brain registering the event, and typically occurs around 300 ms after the infrequent event takes place. This potential is used in our word processor, dubbed the P300 Speller. The P300 speller is described in [8] and was originally proposed by Farwell and Donchin [9]. A modified version of the P300 speller with a reduced frequency of the anticipated event (which increases the P300 potential, enhancing the final accuracy) has been presented in [10].

The P300 speller presents the user with a 6×6 matrix of characters (see Fig. 1). The user's task is to focus his or her attention on the characters of a predefined phrase, one character at a time. All the rows and columns of this matrix are successively and randomly intensified for 100 ms, followed by 75 ms of no intensification. Two out of the 12 intensifications of the rows or columns contain the desired character. The responses evoked by these infrequent stimuli (i.e., the two out of 12 stimuli that contain the desired character) are different from those evoked by the stimuli that do not contain the desired character. The difference in the shape of the response, which is exploited by a classifier, is the working principle behind the P300 speller.

B. Data Acquisition System and Procedure

Our data acquisition system is based on a Neuroscan amplifier called SynAmps2, which has 64 monopolar channels, and four bipolar channels. The bipolar channels are used to monitor eye movement or blinking artifacts during the data collection. The Neuroscan software pipes the data in "server" mode to a TCP/IP port. The data are read by our own software running on a different machine. Our software has two parts. The first part written in C++ collects, records, and processed the data. The

second part (implemented in Java) presents the six by six matrix and the intensification of the desired row or column (the stimulus) to the subjects. The subjects are placed about 3 ft in front of a 15 in liquid crystal display (LCD) panel, where the stimuli are presented. When a particular stimulus is presented, our software sends a corresponding stimulus code to be embedded in the data stream in a time-locked fashion. Using the embedded stimulus code, the data are divided into epochs of 500 ms starting from the time of stimulus presentation.

We have collected data from nine different subjects in separate sessions. All the subjects are healthy male volunteers between 20 and 40 years of age. They consented to participate in the experiment and received no monetary compensation for their contribution.

The data collection procedure has three stages: 1) subject preparation; 2) training data collection; and 3) test data set collection. In total, an experimental session lasts typically about 1 h and 30 min.

During the subject preparation stage, an electro-cap is attached to the subjects and an electrolyte gel is applied to the electrodes to reduce the impedance. The subjects are then briefed on the training and testing procedures. During the training phase of the data collection, the subjects are given a sequence of characters to focus on. Each character is indicated to them by highlighting (intensifying) it on the subjects' LCD display for 4 s. There is a preparatory gap of 2.5 s after the intensification, and the subjects are then shown 10 rounds of 12 visual stimuli. After each round, there is a small pause of half a second before the next round is started.

Once the training data are collected, our SVM classifier is trained. The SVM is then used as an online classifier to collect another round of data. During this test data set collection, the subjects are verbally asked to focus on the characters of a word given by the experimenter. The rows and columns are then randomly flashed. After ten rounds, our software shows the recognized character as a feedback to the subjects. Then there is rest period of 2.5 s before the program moves to the next character in the word. Once the data for one word are collected, the experimenter proceeds to the next word, after ensuring that the subjects are rested and ready to continue. For both training and testing, we use the phrase "THE QUICK BROWN FOX JUMPS OVER LAZY DOG 24 613 8579" for the nine subjects.

C. Signal Processing and Classification

The EEG data are collected at a sampling rate of 250 Hz. At the preprocessing stage, we perform a low-pass filtering of the epoched data, using an optimal cutoff frequency [3]. The data are then down-sampled with a moving-averaging window in order to reduce the data size, thereby making the online classification possible. We use 25 channels out of the 64 available. The manual channel selection is around C3, C4, Cz, CPz, and FCz, in addition to two distant positions P7 and P8. The same 25 channels are used for all the subjects. (An automatic channel selection strategy was explored later [11], but not used in the study reported here.) Subsequent to the manual selection, we apply a principal component analysis (PCA) to further reduce the number of channels, transforming the 25 channel data to 20 channels using the 20 largest eigenvalues. The PCA channel selection also whitens the data in different channels.

During the signal processing stage, ocular artifacts are removed [12], [13] from the data by treating the measured EEG $y(n)$ as a linear superposition of the measured EOG signals $u(n)$ and the real EEG $w(n)$

$$y(n) = \sum_{i=1}^{N} b_i u_i(n) + w(n). \qquad (1)$$

Here, N is the number of sites at which the EOG measurement is done, two in our setup. Since the dynamic range of w is small in comparison to u, we can use least square minimization to compute the propagation constants b_i. The difference model [(2) below] removes the inter-sample correlations of the required EEG signal, $w(n)$

$$y(n) = y(n') + \sum_{i=1}^{N} b_i(u_i(n) - u(n')) + w(n) - w(n') \quad (2)$$

where $n' = n - 1$. In our experiments, we found that (2) outperformed (1) consistently. The artifacts are removed online using (2) throughout the studies reported here.

SVM [14] is then used for classification. SVM is a powerful approach for pattern recognition. It provides a very good discriminative solution with good generalization characteristics. It is suitable for our case, where we normally only have a relatively small amount of training data due to the nontrivial efforts needed for data collection.

The input to the SVM classifier is a data vector formed by concatenating vectors from all channels as follows:

$$x = [x(1)^T, \ldots, x(N)^T, \Delta x(1)^T, \ldots, \Delta x(N)^T]^T \quad (3)$$

where T denotes matrix transpose, and N is the number of samples, and $x(n)$ is a K-channel vector at time instance n, $x \in R^{2 \times N \times K}$. Δx is a dynamic feature defined as

$$\Delta x(n) \approx A \sum_{m=-1}^{W} m\, x(n+m) \qquad (4)$$

where A is a normalization constant. Δx is a least square estimate of the time derivative of the EEG data. It supplies extra information to the classification, improving the accuracy [10].

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

$$f(x) = \text{sgn}\left(\sum_{i=1}^{m} y_i \alpha_i k(x, x_i) + b\right) \qquad (5)$$

where $x \in R^{2 \times N \times K}$ are the test data, $x_i \in R^{2 \times N \times K}$ are the training data support vectors, $y_i \in \{-1, 1\}$ are the class labels, and $k(\cdot)$ is the kernel function. A Gaussian kernel (reasonable choice for EEG signal) function is used. Instead of using (5) to directly classify each epoch, we sum up the SVM margins for all the data for the same character and then make a decision for its row and column separately. For example, after a number of
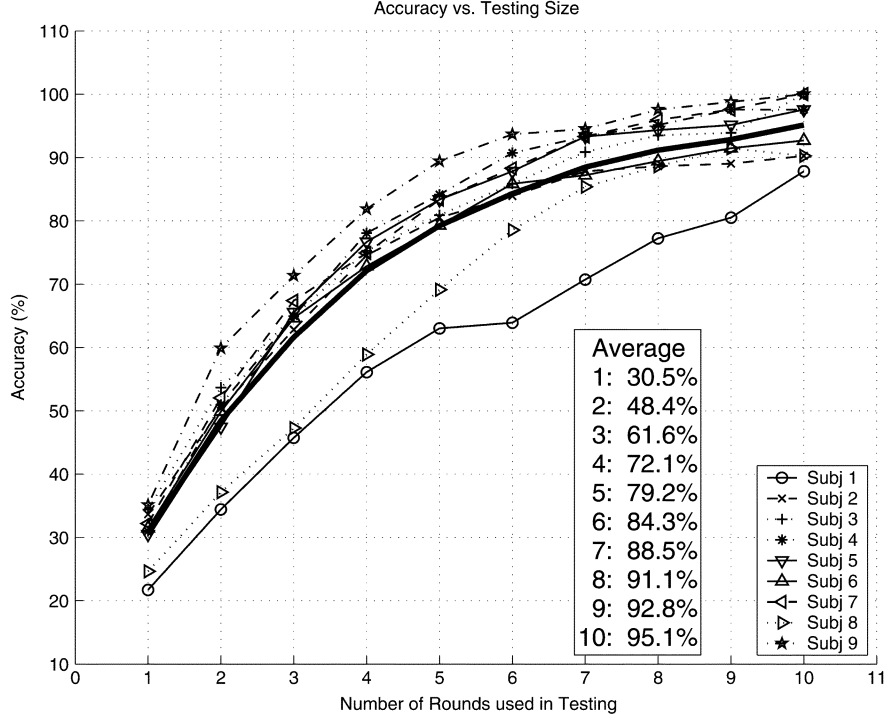
Fig. 2. Dependence of character accuracy on the number of rounds used in testing. Thick line is the average.

rounds of stimuli are presented, we decide the most probable column and row by the following decision functions:

$$ n_{c|r} = \operatorname*{arg\,max}_{n=1,\dots,N_{c|r}} \left\{ \sum_{j=1}^{R} \left[\sum_{i=1}^{m} y_i \alpha_i k \left(x_{c|r,j}^{(n)}, x_i + b \right) \right] \right\} $$

(6)

where $N_{c|r}$ is the number of columns (or rows) in the alphanumeric matrix (both are six in our case), R the number of rounds, and $x_{c|r,i}$ is the test data for columns (or rows). Once a column index and a row index are determined, a character is recognized. Hence, we do not average the signal in the data space but in the classification score space, thus ensuring the trained SVM classifier can be used online for the P300 speller with an arbitrary number of rounds.

D. Reducing Learning Time Requirement

In order to make the P300 system easy to use, the learning (or training) time has to be reduced to a minimum. In our paradigm, all the rows and columns of the spelling matrix (Fig. 1) are randomly flashed once per round; i.e., there are 12 flashes per round. And the entire round is repeated ten times. Thus, there are ten rounds per character. Since we use 41 characters for training, the training phase of the data collection takes about 20 min. Though modest, we would like to make it even smaller to improve the usability of the system. The training time required depends directly on the number of characters and number of rounds used. Since we have the training data with 41 characters × 10 rounds, we can study the dependence of the accuracy

TABLE I
INTER-SUBJECT ACCURACIES. EACH CELL IS THE CHARACTER ACCURACY OF THE SUBJECT (ROW NUMBER) WHILE TESTED AGAINST THE MODEL OF A DIFFERENT SUBJECT (COLUMN NUMBER)

Subj.	1	2	3	4	5	6	7	8	9
1	88	29	24	51	39	24	20	17	46
2	37	90	10	24	17	2	15	7	32
3	54	41	100	68	32	54	39	56	63
4	32	7	22	98	32	39	15	20	20
5	27	17	7	20	98	5	12	17	20
6	39	2	56	51	56	93	34	46	22
7	12	7	7	2	5	22	100	34	24
8	20	7	5	12	32	15	34	90	10
9	59	71	56	59	20	46	63	41	100

with an SVM model created with a subset of the training data. We perform this study in two modes:

1) as a function of the number of rounds for each character, while using all 41 characters;
2) as a function of the number of characters used in the training data set using all ten rounds.

The analysis procedure for both these modes is similar. We retrain the SVM with a subset of the total available data and create a model. We then use that model to analyze the testing data. The results are presented and discussed below.

III. RESULTS

We have characterized the data several different ways. First, we examined spelling accuracy, for each subject, as a function of the number of rounds of data collected. Second, we studied the specificity of the SVM model to each subject as an inter-subject dependency. Finally, we optimized the amount of training needed to make a high accuracy word speller system.
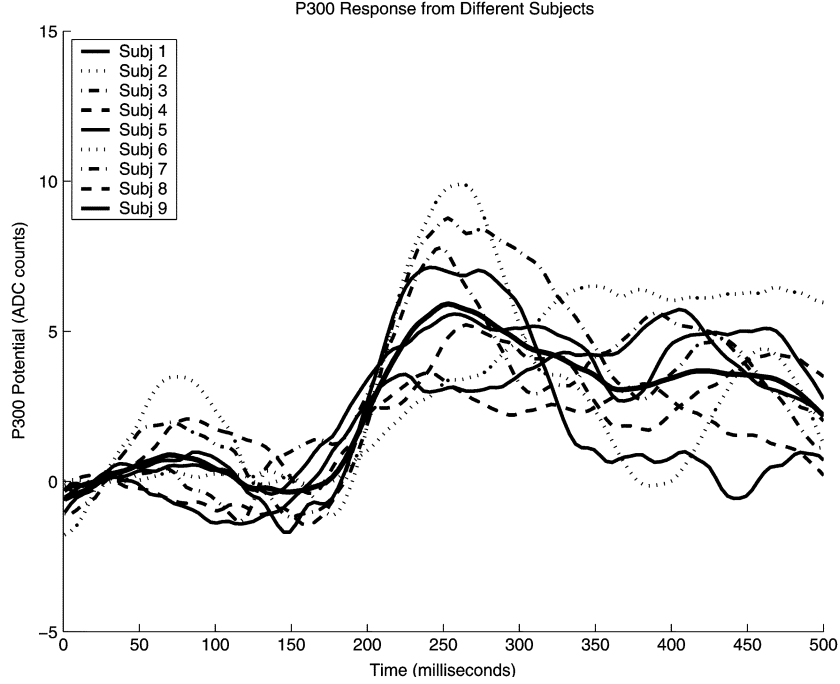
Fig. 3. P300 potential from the nine subjects. P300 is measured at Cz with the reference set to the average of P7 and P8 potentials. Baseline correction from the first 50 ms is applied. X axis is the time from the stimulus presentation. Each curve is the average of 820 stimulations (41 characters × 10 rounds × 2 on-target flashes).

A. Word Speller Accuracy

With a view to minimizing the time required to spell a character, we studied the variation of character accuracy as a function of the number of rounds. This variation is shown in Fig. 2. In this figure, the *x* axis is the number of rounds, and the Y axis is the accuracy attained if the classification is done using only as many rounds as in the *x* axis. The final accuracies reported in this article correspond to the tenth round (i.e., the points in the figure at $X = 10$).

Fig. 2 helps us determine the number of rounds where the accuracy saturates; repeating the data collection beyond this point serves no purpose. The average accuracy is about 90% by the seventh round. It then smoothly increases to about 95% by the tenth round. Thus, a high accuracy word speller system using this paradigm requires all ten rounds.

B. Inter-Subject Dependence

We studied the variations of P300 signal between different subjects. We tested the data of each subject against the models created from the others. The results are shown as a matrix in Table I. Here, we used all ten rounds. The diagonal elements are accuracies when a subject is tested against his own model. The low values of the off-diagonal elements indicate the high degree of subject specificity in the data. This finding raises the possibility that the P300 signals may be significantly different between different subjects. In Fig. 3, we have plotted the P300 potential from the nine subjects. It shows differences in the latency, amplitude, or even the shape of P300 potentials, which may have their origin in the electrode locations on the scalp in addition to the inherent subject-to-subject variations.

C. Minimum Training Required

Fig. 4 shows the accuracy as a function of the number of rounds used in training the SVM. It shows that the average accuracy plateaus at the seventh round, immediately suggesting that we can reduce the duration of the training session by 30%. Note that during this study, we used all 41 characters in the training set.

Fig. 5 shows the accuracy as a function of the number of characters used in the training sample. In this study, the training data set was divided into subsets consisting of one or more words of the training phrase "THE QUICK BROWN FOX JUMPS OVER LAZY DOG 24 613 8579". Thus, we measured the accuracy using 3, 8, 13, 16, 21, 25, 29, 32, 37, and 41 characters. The average accuracy saturates at about 25 characters. Increasing the training data size beyond 25 characters, therefore, serves no purpose. Using only 25 characters for training results in a 40% reduction in the training time required.

These results show that the training time requirement can be reduced by about 58% if there is no correlation between the number of rounds and the number of characters used in the training dataset. Such an assumption of the lack of correlation is unreasonable, as Table II shows. The effect of reducing both the number of rounds (by 30%) and the number of characters (by 40%) together is a decrease of about 5% in the average accuracy.

IV. DISCUSSION AND CONCLUSION

We have presented our studies on a word speller system based on the P300 evoked potential, with an average accuracy of about 95%. The time taken for each character is about 22 s. Although our training data collection lasted about 20 min, we showed that it can be reduced to less than 10 min with only about a 5% drop
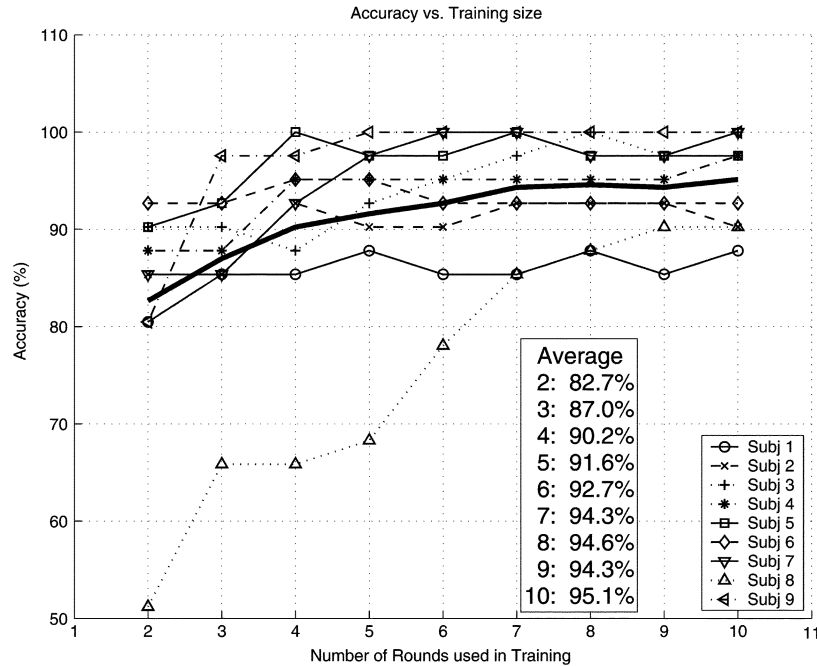
Fig. 4. Dependence of character accuracy on the number of rounds used in training. Thick line is the average of the nine accuracies plotted. Due to a technical difficulty involving the dimensionality of the data matrix, we could not do the training with just one round. Hence, the missing data point for the single round.
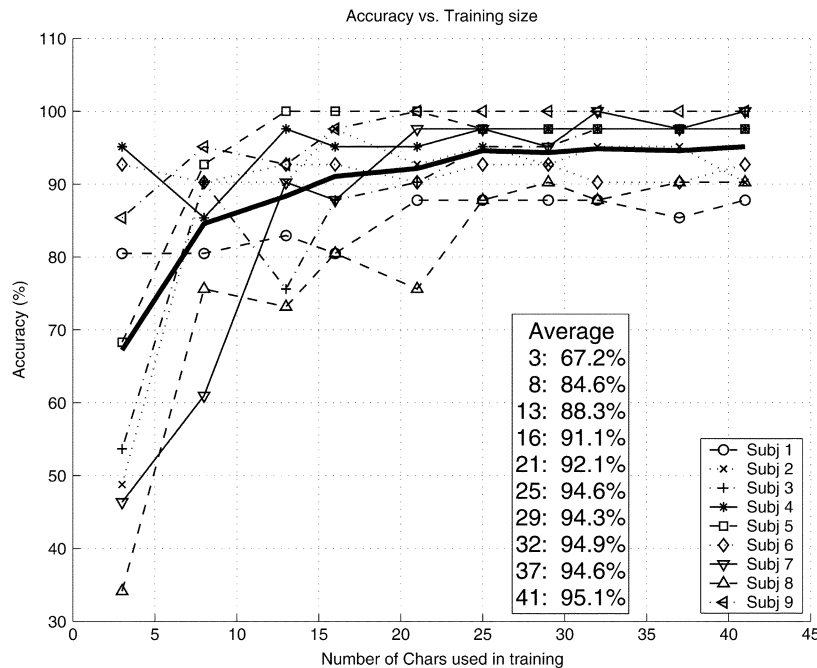

Fig. 5. Dependence of character accuracy on the number of characters used in training. Thick line is the average of the individual accuracies.

in the accuracy. We made significant improvements in the accuracy and speed by employing powerful machine learning algorithms for classification and developing a new dynamic feature. These novel approaches decrease the error rate significantly (by a factor of two or three) over the standard methods such as P300 peak finding and area methods. This P300 speller can be used in an online mode. Although our studies were done on healthy subjects, there is a chance that BCI systems such as the one presented in this paper may some day provide potentially the only communication channel for severely disabled

people who are otherwise unable to articulate their thoughts and needs.

TABLE II
ACCURACIES WHEN THE SVM CLASSIFIER IS TRAINED WITH VARIOUS AMOUNT OF TRAINING DATA. THE PERCENTAGE OF THE TRAINING DATA USED IS SHOWN IN THE PARENTHESES

Characters → Rounds ↓	21 (51%)	25 (61%)
6 (60%)	88.4%	89.0%
7 (70%)	89.1%	90.2%

Acknowledgment

The authors would like to thank S. Ranganatha, X. Zhu, W. Xu, and Q. Koh for their help with the data collection.

References

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," in *Clinical Neurophys.*, vol. 113, 2002, pp. 767–791.

[2] J. R. Wolpaw, "Brain–Computer interfaces (BCI's) for communication and control: Current status," *2nd Int. BCI Workshop Training Course*, pp. 29–32, 2004.

[3] C. Guan, X. Zhu, S. Ranganatha, M. Thulasidas, and J. Wu, "Robust classification of event-related potential for brain-computer interface," in *Proc. 2nd Int. Conf. Adv. Medical Signal Info. Processing (MEDSIP)*, Sep. 2004, pp. 321–326.

[4] K. R. Müller, M. Krauledat, G. Dornhege, G. Curio, and B. Blankertz, "Machine learning techniques for brain-computer interfaces," *2nd Int. BCI Workshop Training Course*, pp. 11–22, 2004.

[5] P. Meinicke, M. Kaper, F. Hoppe, M. Heumann, and H. Ritter, "Improving transfer rates in brain computer interfacing: A case study," in *Proc. NIPS*, Canada, 2002, pp. 1107–1114.

[6] M. Kaper, P. Meinicke, U. Grossekathoefer, T. Lingner, and H. Ritter, "BCI competition 2003—Dataset IIb: Support vector machines for the P300 speller paradigm," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1073–1076, Jun. 2004.

[7] M. Kaper and H. Ritter, "Generalizing to new subjects in brain-computer interfacing," in *Proc. 26th IEEE EMBS Annu. Int. Conf. (EMBC)*, Sep. 2004, pp. 4363–4366.

[8] E. Donchin, K. M. Spencer, and R. Wijesinghe, "The mental prosthesis: Assessing the speed of a P300-based brain-computer interface," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 2, pp. 174–179, Jun. 2000.

[9] L. Farwell and E. Donchin, "Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalogr. Clin. Neurophys.*, vol. 70, pp. 510–523, 1988.

[10] C. Guan, M. Thulasidas, and J. Wu, "High performance P300 speller for brain-computer interface," in *Proc. IEEE Int. Workshop Biomed. Circuits Syst.*, Dec. 1–3, 2004.

[11] M. Thulasidas and C. Guan, "Optimization of BCI speller based on P300 potential," in *Proc. 27th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBS)*, Sep. 2005.

[12] M. Thulasidas, C. Guan, R. Ranganatha, J. Wu, X. Zhu, and W. Xu, "Effect of ocular artifact removal in brain computer interface accuracy," in *Proc. 26th IEEE EMBS Annu. Int. Conf. (EMBC)*, vol. 2, Sep. 2004, pp. 4385–4388.

[13] S. M. Haas *et al.*, "EEG ocular artifact removal through ARMAX model system identification using extended least squares," in *Commun. Inf. Sys.*, vol. 3, Jun. 2003, pp. 19–40.

[14] V. N. Vapnik, *Statistical Learning Theory*. New York: Wiley, 1998.

Manoj Thulasidas (M'03) received the undergraduate degree from the Indian Institute of Technology, Madras in 1987. He then joined Syracuse University, Syracuse, NY, as a graduate student, working with the CLEO collaboration at Cornell University, Ithaca, NY, during 1990–1992. He received the Ph.D. degree from Syracuse University in 1993.

After receiving his Ph.D., he moved to Marseilles, France and continued his fundamental research with the ALEPH collaboration at CERN, Geneva. He has over 190 journal publications from his ten-year career as a Research Scientist in the field of high energy physics. He then joined the Kent Ridge Digital Laboratories (later renamed the Institute for Infocomm Research), Singapore, in 1998, to study and develop various human body-based measurements and systems. His work at this institute resulted in four invention disclosures, two of which have been filed for patents, and numerous academic papers. He has also authored a book on the role of perception and cognition in physics.
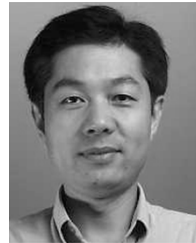
Cuntai Guan (S'91–M'92–SM'03) received the Ph.D. degree in electrical and electronics engineering from Southeast University, Nanjing, China, in 1993.

He has many years of experience in various academic, research, and industrial organizations. He has worked as Associate Professor, Visiting Scientist, Research Manager, and Research and Development Director. His current research focuses on investigation and development of an effective framework and statistical learning algorithms for the analysis and classification of brain signals.
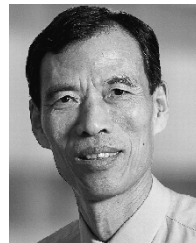
Wu Jiankang received the B.Sc. degree from the University of Science and Technology of China, Hefei, and the Ph.D. degree from Tokyo University, Tokyo, Japan.

Prior to joining the Institute for Infocomm Research, Singapore, in 1992, he was a Full Professor at the University of Science and Technology of China. He also worked in universities in the U.S., U.K., Germany, France, and Japan. He is the author of 18 patents, 60 journal publications, and five books.

Dr. Jiankang has received nine distinguished awards from China and the Chinese Academy of Science.