ASYNCHRONOUS P300 BCI: SSVEP-BASED CONTROL STATE DETECTION

Rajesh C. Panicker[1], Sadasivan Puthusserypady[2], Ananda P Pryana [1] and Ying Sun[1]

Department of Electrical and Computer Engineering[1],
National University of Singapore, 4 Engineering
Drive 3, Singapore 117576.

Department of Electrical Engineering[2],
Technical University of Denmark, Lyngby, Denmark
phone: 45-45253652, fax: 45-45880117, email:
spu@elektro.dtu.dk.

ABSTRACT

An asynchronous hybrid brain-computer interface (BCI) system combining the P300 and steady-state visually evoked potentials (SSVEP) paradigms is introduced. A P300 base system is used for information transfer, and is augmented to include SSVEP for control state detection. The proposed system has been validated through off-line and online experiments. It is shown to achieve fast and accurate control state detection without significantly compromising the performance. For the two subjects who participated in the online experiments, the system achieved an average data transfer rate of 20.13 bits/min, with control state classification accuracy of more than 97%.

1. INTRODUCTION

Brain-computer interface (BCI) is a system which can be used for direct communication with a computer system, without reliance on the neuromuscular pathways. The brain activity patterns are detected and translated into control signals for a computer or a prosthetic device. The typical patterns in the electroencephalogram (EEG, the electrical activity of brain measured on scalp) exploited by BCI systems are evoked potentials (EPs), event related potentials (ERP), motor imagery, various band rhythms etc [1] etc. Two widely used potentials in BCI research, found in EEG are P300 ERP and steady-state visually evoked potentials (SSVEP). P300 can be detected as a positive peak predominant at centro-parietal region, about 300ms after the presentation of a rare, task-relevant stimulus [2]. Its amplitude and latency varies between individuals and even within the same individual over time, and also depending on the type of stimulus [3]. SSVEP is another widely used potential, produced by the brain in response to repetitive periodic visual stimulus. When the subject focuses on a visual stimulus steadily flickering at a certain frequency in the range of 3-75 Hz, the brain produces a detectable signal of the same frequency and its harmonics [4]. By presenting several stimuli, each flickering at different frequencies, researchers have developed robust BCI systems capable of reaching very high average information transfer speed of up to 68 bits/minute [5].

A very desirable feature of a practical BCI is the capability for asynchronous operation. The BCI system should be able to detect if the user is intending to input a command at all, instead of expecting the user to issue one command at every fixed interval throughout the period of operation. Hence, asynchronous BCI has become an active field of research and encouraging results have been reported [6]. However, very few P300-based asynchronous systems have been reported. Zhang et al. developed an asynchronous P300 speller which is able to communicate at an average of 15 bits/min, and a false positive rate of 0.7 events/min [7]. They achieved asynchronous control by setting a threshold for the likelihood derived from a probabilistic model of P300 classifier scores. However, given the high inter and intra-subject variability of P300 response, the model parameters may not hold for extended periods of time, and maintaining accuracy without frequent re-training/updating model parameters would be challenging. An alternate method to develop an asynchronous system is proposed in this paper. The idea is to use different EPs for control state detection and information transfer. In this paper, we propose a P300 based system, with SSVEP providing the control state information.

The base system described here utilizes the P300 ERP. The advantages of P300 is its suitability for a wide spectrum of users including disabled patients [8], relaxed requirement of visual attention and relative ease of detection, and reasonably good information transfer rates. Though SSVEP based systems are generally faster than P300 based systems, they suffer from several drawbacks such as the requirement of accurate control of eye-muscles [9], precise and fast hardware, and unsuitability for people with epilepsy. Moreover, if low frequency stimuli are used, prolonged use of the system is very tiring whereas high frequency SSVEP response is weaker and harder to detect accurately. Here, instead of basing the complete system on SSVEP, we utilize it just for control state detection, with P300 as the main BCI paradigm. Following the terminology used in [7], the EEG data associated with the flashing of one button, and that associated with one complete cycle of flashings are called *epoch* and *round* respectively in this paper. Also, the state in which the user is actively giving an input is called *control state*, and *non-control state* otherwise. The proposed method and the experimental setup is described in section 2. Section 3 details the data analysis and section 4 describes results for off-line and online experiments. The paper is concluded with some remarks in section 5.

2. HYBRID P300-SSVEP SYSTEM

SSVEP is an ideal candidate to be used in conjunction with the P300 ERP, for several reasons. Both are well documented to be reliably evoked in virtually everyone without any need of prior training. Unlike a pure SSVEP based system, P300 BCI does not require precise control over eye muscles, making it suitable for severely disabled patients. The visual stimulus required to elicit SSVEP can be added to the existing P300 stimuli with relative ease, as both are usually evoked by a visual stimuli (P300 can also be evoked by other stimuli, but visual P300 is dominant in BCI research). Our experiments show that both the signals can be elicited at the same

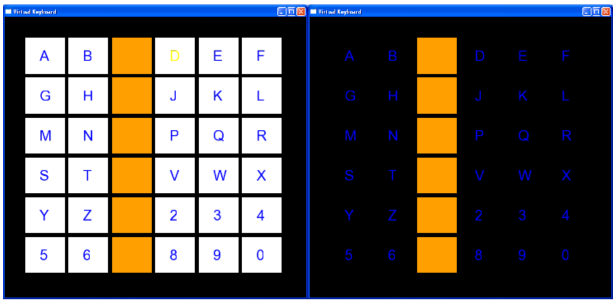


Figure 1: The screen will flicker from white to black rapidly. The row/column flashes are carried out just as in a standard P300 speller

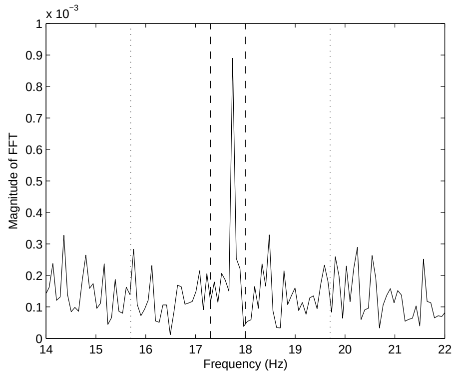


Figure 2: The peak picking algorithm. The objective function is ratio of power in the band enclosed by the thick lines as against the power in the band enclosed by the thin lines.

time in an individual, without greatly compromising the detection accuracy of either.

The experimental setup makes use of a 24 channel amplifier from ANT-Neuro, with a sampling rate of 256Hz. EEG from 9 channels of the standard 10-20 system [10] - Cz, C1, C2, Pz, P1, P2, Oz, O1 and O2, were recorded. The data recording is controlled from a multi-threaded program implemented in Visual C++ through the ActiveX control. Another thread handles the interface, which is the speller paradigm [11] implemented using SFML - a multimedia library providing accelerated graphics using OpenGL as back-end. The speller consists of 36 characters, arranged as a 6×6 matrix with characters A-Z and 0-9. The rows and columns are highlighted in a random order such that all rows and columns are highlighted once in every round. When the user is concentrating on one particular character, a P300 is elicited when either the row or column containing the character is flashed. Precise timing is ensured by recording all time stamps from the same timer. The processing is done in real time using MATLAB. A third thread waits for decisions and passes it to the display interface. In our system, the whole display is set to flicker at the desired frequency to elicit the SSVEP while the normal highlighting of rows and columns is done as usual for the P300 based interface. Figure 1 shows the two alternating states. When the user is gazing at the screen, it can be assumed that he/she wishes to input a command, which will manifest as the elicitation of SSVEP. With such an interface, the user would be able to naturally elicit both potentials without having to pay a split attention. Thus the ability to evaluate both the control state and the actual decision at all times is a significant advantage.

Since only one frequency is used in the SSVEP detection, the precise attention requirement of SSVEP is eliminated. Hence, the task is reduced to the detection of any SSVEP near the frequency of interest, as opposed to precisely detecting one among several frequencies. Thus, the need for a dedicated hardware capable of creating very precise stimuli of various frequencies is also eliminated; cheap and simple displays would be sufficient. By choosing the frequency to be outside the usual range for P300 (i.e. above 12Hz), the two signals could be separated by simple bandpass filtering and thus there would be no reduction of accuracy in the classification process.

Preliminary experiments were conducted to explore the best flicker frequency to use in the subsequent experiments and in the online experiments. A stimulus frequency (f_{st}) of around 18Hz was chosen due to the following reasons: (i) a frequency lower than 15Hz interferes with P300 signal, adversely affecting the classification accuracy, (ii) there is a trade off between user comfort (higher frequency is preferred) and SSVEP amplitude and hence detection accuracy (lower frequency is better), and (iii) higher frequency is very demanding on hardware.

3. DATA ANALYSIS

3.1 SSVEP Detection

SSVEP is usually very precise about the stimulus frequency. Gao et al. reported the possibility of distinguishing two stimulus with frequency difference of just 0.2Hz [5]. Detection of SSVEP is usually done by a simple thresholding of the amplitude of signal's FFT. More advanced methods like canonical correlation analysis (CCA) have been proposed and are shown to give excellent SSVEP classification [12]. Various techniques for enhanced detection of SSVEP can be found in [13–15].

Unlike the techniques mentioned above, the detection task in our system is less demanding on frequency precision, as the presence or absence of SSVEP is all that is required to be estimated. Therefore, in this control state detection scheme, all other peaks not located around the target frequency can be safely assumed to be due to noise and ignored. Simple thresholding of band-power would not work due to high variability of EEG signals and the presence of a salient peak at the target frequency needs to be ascertained. Hence, the mean power in a wider range of frequencies is used as a benchmark for comparison. Figure 2 shows a sample FFT result of an epoch in which the user used the hybrid system with the screen flickering at around 18Hz. The relative mean power spectral density (PSD) of frequency bins in the narrow range $f_{st} \pm f_n$ Hz as compared to the mean PSD in the wider range $f_{st} \pm f_w$ Hz is the metric chosen for detection. Thus, an objective function can be defined as

$$J(f_{st}) = \frac{[S(f)]_{f_{st} \pm f_n} - [S(f)]_{f_{st} \pm f_w}}{[S(f)]_{f_{st} \pm f_w}} \quad (1)$$

where $[S(f)]_{f_{st} \pm f_n}$ is the mean PSD in the narrow range and $[S(f)]_{f_{st} \pm f_w}$ is the mean PSD in the wider range. In our experiments, f_n is chosen to be 0.3Hz, and f_w is chosen to be 2Hz. The value of $J(f_{st})$ could then be compared with a threshold to detect SSVEP, which in turn indicates user's desire to input a command. The frequency sensitivity of the algorithm could be tuned by setting the ranges. The threshold controls the balance between true positive rate (TPR) and false positive rate (FPR), the setting of which depends on the specific application. Channel selection was done based on the inspection of power spectral densities of the data from various channels at the frequency of interest.

3.2 P300 Classification

The collected data is bandpass (zero-phase) filtered between 0.5 Hz and 12 Hz using a Butterworth filter of order 3. To reduce the feature size, it is down-sampled to 32Hz, and data for a duration of 0.7 seconds from the start of the stimulus is considered to belong to that particular epoch. Reliable detection of P300 usually requires several rounds. The optimum number of rounds to be chosen is a trade-off between classification accuracy and the information transfer rate (ITR, computed based on the suggestion by Wolpaw *et al.* [16,17]), and varies from person to person. The number of rounds used for the detection of a character is fixed to be 5, as it was found to be giving a near-perfect accuracy in our preliminary experiments.

3.3 FLDA

In FLDA, the data is projected to a lower dimension such that the projected means of the classes are far apart, while the spread of projected data is small. This can be realized by optimizing a cost function related to within-class matrix ($\mathbf{S}_w$) and between-class matrix ($\mathbf{S}_b$), which are defined as

$$\mathbf{S}_w = \sum_{k=1}^{n^c} \sum_{\mathbf{x}_j \in c_k} (\mathbf{x}_j - \mathbf{m}_k)(\mathbf{x}_j - \mathbf{m}_k)^T \tag{2}$$

$$\mathbf{S}_b = \sum_{k=1}^{n^c} n_k^c (\mathbf{m}_k - \mathbf{m})(\mathbf{m}_k - \mathbf{m})^T \tag{3}$$

where $\mathbf{x}_j$; $j = 1, 2, \ldots, n^t$ are training data vectors, $\mathbf{x}_j \in c_k$ denotes all $\mathbf{x}_j$ belonging to the k^{th} class, $\mathbf{m}_k$ is the mean of samples belonging to the k^{th} class, $\mathbf{m}$ is the global mean, n^c is the number of classes (n^c=2 in our classification, denoting either the presence or the absence of P300), and n_k^c is the number of samples in the k^{th} class. Given the pattern matrix $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{n^t}]$ and the corresponding label vector $\mathbf{y} = [y_1, y_2, \ldots, y_{n^t}]$, the problem is to find a projection vector $\mathbf{w} = [w_1, w_2, \ldots, w_d]^T$ such that the projection $\mathbf{y} = \mathbf{w}^T\mathbf{X}$ maximizes the criterion function

$$J_p(\mathbf{w}) = \frac{\det(\mathbf{w}^T \mathbf{S}_b \mathbf{w})}{\det(\mathbf{w}^T \mathbf{S}_w \mathbf{w})}. \tag{4}$$

The solution [18] is to choose $\mathbf{w}$ satisfying the eigen equation

$$\mathbf{S}_w^{-1} \mathbf{S}_b \mathbf{w} = \lambda \mathbf{w}, \tag{5}$$

if $\mathbf{S}_w^{-1}$ exists, λ being the only non-zero eigenvalue of $\mathbf{S}_w^{-1}\mathbf{S}_b$. Once $\mathbf{w}$ is estimated, the classifier design is complete and the output for a single feature vector is

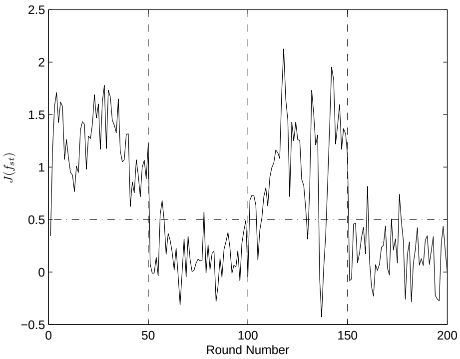
$$y_j = \mathbf{w}^T \mathbf{x}_j. \tag{6}$$

In each round, the scores for all rows and columns are calculated using Eq.(6). The estimated target is the symbol at the intersection of the row and the column having the maximum of the averaged scores.

4. RESULTS AND DISCUSSION

4.1 Off-line Experiments

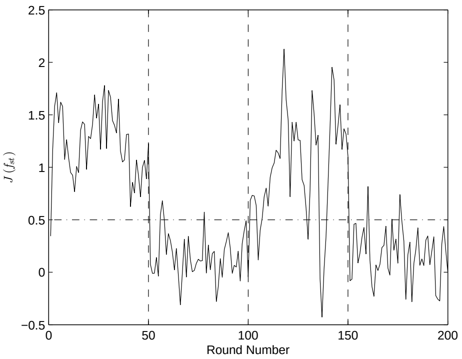
To evaluate the performance of the proposed scheme, off-line experiments were conducted on three healthy subjects aged 22-27 (two males and one female). For training a P300 classifier, EEG for 300 rounds of stimuli were recorded, with the target character highlighted during the session. Each subject performed an experiment of 40 characters with an inter-stimulus interval (ISI) of 200ms. Subjects are in control state for the first 10 characters (with 5 rounds per character), and in non-control state for the next 10 characters and so on. In control state, the subject is instructed to count the number of times the target character is highlighted. The subject is instructed to do a mental task (multiplication) and to relax with eyes closed for alternate non-control states. The P300 detection accuracy was comparable to that obtained in normal experiments - out of the 20 characters that the subject focused on, 20, 19 and 18 characters were correctly classified for the subjects 1, 2 and 3 respectively. The spectrum of the first 20 characters for subject 1 is shown in Fig. 3, which clearly shows that with a full block of data, distinguishing between control and non-control states can be easily done.



(a) subject 1

Figure 3: FFT of the first 20 characters for subject 1. The characters 1-10 are in control state

From the experimental data it was found that the data from just one round is not always sufficient for reliable control state detection. However, it is not necessary to follow P300's trial demarcation rigidly in this case, and it is possible to obtain more data per round without lengthening ISI and sacrificing bit rates simply by allowing some overlap of data between rounds. It is justified if we assume that the user would have been focusing on the screen for at least a few seconds before the onset of the stimulus. By extending the data for a round to include the data from 2 seconds before the start of the P300 stimulus, the classification accuracy was found to be better. A sample result obtained by evaluating the objective function using Eq.(1) for each round with the extension as mentioned above is given in Fig.4. The vertical (dashed) lines indicate a change in control state, and as expected, the non-control rounds have values around zero. The horizontal dashed line is the threshold setting used. If the goal is to maximize the classification accuracy, the threshold can be found using exhaustive search of a portion of the data. For example, the optimum threshold obtained for subject 1 is 0.58 and 188 out of 200 rounds were successfully classified (94%).



(a) subject 1

Figure 4: $J(f_{st})$ for the subject 1

To evaluate the performance of the system for various thresholds, the area under curve (AUC) of the receiver-operating characteristic (ROC) was computed. The ROC is given in Fig.5, and the AUCs are summarized in Table.1. Given that AUC for a random classification is 0.5, the detection capability of the system is very good. The classification accuracy (CA) of the speller, the corresponding ITR as well as control state detection accuracies (CD) for the subjects are also given therein. Voting of classifier labels within a block was used for the calculation of CD. For example, if blocks of 5 rounds are considered for the detection of one character; as long as at least 3 rounds are determined to be in control state, the character is deemed valid.

4.2 Online Experiments

The online experiment is implemented as semi-asynchronous. The BCI system is still operated in a discrete, predefined blocks of rounds. In this experiment, 5 rounds per block and an ISI of 200ms were used. The signal data was overlapped for the purpose of SSVEP detection. Once stimuli for one block is finished, the system will halt until a decision has been made, and a new block will start. In addition to detecting target character using P300 in each round, the presence of SSVEP is checked for to validate the detection. As long as SSVEP is detected in at least 3 out of 5 rounds, the subject is deemed to be in control state. P300 classification is employed only when control state has been established.

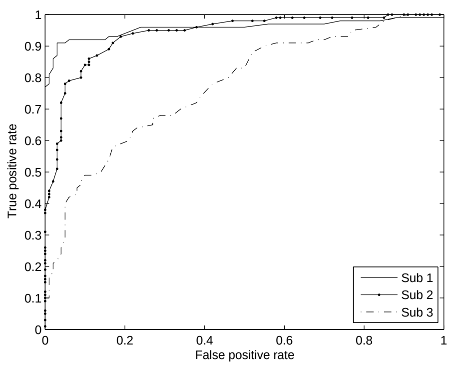
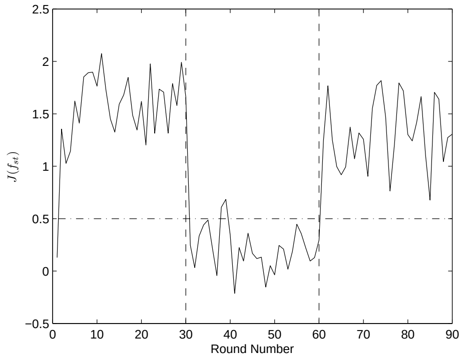


Figure 5: ROC curve for the subjects



(a) subject 1

Figure 6: $J(f_{st})$ for the subject 1 in the online experiment

Table 2: Detection results for the online experiment. CS and NCS are the average SSVEP detections for blocks of 5 rounds, when the subject is in control state and non-control state respectively. CD is the block-wise correct detection of control state.

Subject	CS	NCS	CD	CA	ITR (bits/min)
Sub 1	4.88	0.12	96.30	86.11	17.89
Sub 2	4.78	0.15	98.15	97.22	22.36
Average	4.83	0.135	97.225	91.665	20.13

The two best performing subjects (subjects 1 and 2) from the off-line experiments participated in the online experiments. These subjects performed three runs of 18 characters each. Each character is determined once 5 rounds have been presented, and the character is determined to be null if control state is not detected. In each run, the subject focused on the first 6 characters, gazed away for the next 6, and focused again on the last 6. Thus, there would be 54 blocks of 5 rounds each, 36 of which are in control state. The threshold value was set to be 0.5. $J(f_{st})$ for the subject 1 in the experiment can be found in Fig.6. In blocks of 5, the average detections for control states and non-controls states were 4.88 and 0.12 respectively for subject 1 and 4.78 and 0.15 respectively for subject 2. The corresponding control state detection accuracies were 96.30 and 98.15 respectively; which shows that the control state detection using SSVEP is very robust. It was noted that the accuracy is lower when focusing on the last column of the display, likely due to the reduced visual attention to SSVEP. Based on the P300 detection accuracy, the system is capable of information transfer at 17.89 bits/min and 22.36 bits/min respectively for the subjects if he/she is continuously in control state. The results are summarized in Table. 2. This is comparable to the results obtained without the SSVEP and it was observed that the addition of SSVEP for control state detection does not affect the accuracy of P300 classification significantly.

Table 1: Detection results for the off-line experiment

Rounds/char		2			3			5		
Subject	AUC	CA (%)	ITR (bits/min)	CD (%)	CA (%)	ITR (bits/min)	CD (%)	CA (%)	ITR (bits/min)	CD (%)
Sub 1	0.958	65	25.24	85	95	33.86	95	100	23.86	97.5
Sub 2	0.928	90	47.87	72.5	95	33.86	97.5	95	21.36	95
Sub 3	0.751	60	22.21	50	70	20.12	65	90	19.32	80
Average	0.879	71.67	31.77	69.17	86.67	29.28	85.83	95.00	21.51	90.83

5. CONCLUSIONS

An asynchronous hybrid BCI system combining two different paradigms has been realized. This system takes advantage of the ease of elicitation of the SSVEP, and flexibility of P300 such that the system has efficient operation and reliable control state detection. The system achieved an ITR of 20.13 bits/min, with a control state classification accuracy of more than 97%. Higher frequency flicker could be used to increase the user's comfort when using the system [19]. This would necessitate sophisticated hardware and detection algorithms. It should be noted that the performance of the system for one subject was relatively low as compared to the other two. A detailed analysis with large pool of subjects would be vital to throw more light into the subjectivity of the system. As the system involves flickering stimuli, a detailed study of factors affecting user comfort, and effects of habituation needs to be done.

REFERENCES

[1] G. Dornhege, J. D. R. Millan, T. Hinterberger, D. J. Mcfarland, K.-R. Muller, and T. J. Sejnowski, *Toward Brain-Computer Interfacing*. MIT Press, 2007.

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

[3] B. Allison and J. Pineda, "ERPs evoked by different matrix sizes: implications for a brain computer interface (BCI) system," *IEEE Trans. Neural Sys. Rehab. Eng.*, vol. 11, no. 2, pp. 110–113, June 2003.

[4] F. Beverina, G. Palmas, S. Silvoni, F. Piccione, and S. Giove, "User adaptive BCIs: SSVEP and P300 based interfaces," *PsychNology Journal*, vol. 1, no. 4, pp. 331–354, 2003.

[5] X. Gao, D. Xu, M. Cheng, and S. Gao, "A BCI-based environmental controller for the motion-disabled," *IEEE Trans. Neural Sys. Rehab. Eng.*, vol. 11, no. 2, pp. 137–140, 2003.

[6] G. Townsend, B. Graimann, and G. Pfurtscheller, "Continuous EEG classification during motor imagery-simulation of an asynchronous BCI," *IEEE Trans. Neural Sys. Rehab. Eng.*, vol. 12, no. 2, pp. 258–265, June 2004.

[7] H. Zhang, C. Guan, and C. Wang, "Asynchronous P300-based brain-computer interfaces: A computational approach with statistical models," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 6, pp. 1754–1763, 2008.

[8] F. Nijboer, E. Sellers, J. Mellinger, M. Jordan, T. Matuz, A. Furdea, S. Halder, U. Mochty, D. Krusienski, T. Vaughan *et al.*, "A P300-based brain-computer interface for people with amyotrophic lateral sclerosis," *Clin. Neurophysiol.*, vol. 119, no. 8, pp. 1909–1916, 2008.

[9] J. Wolpaw, N. Birbaumer, D. McFarland, G. Pfurtscheller, and T. Vaughan, "Brain-computer interfaces for communication and control," *Clin. Neurophysiol.*, vol. 113, no. 6, pp. 767–791, 2002.

[10] G. H. Klem, H. O. Lüders, H. H. Jasper, and C. Elger, "The ten-twenty electrode system of the international federation. the international federation of clinical neurophysiology." *Electroenceph. Clin. Neurophysiol.*, vol. 52, pp. 3–6, 1999.

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

[12] Z. Lin, C. Zhang, W. Wu, and X. Gao, "Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs," *IEEE Tran. Biomed. Eng.*, vol. 54, no. 6, pp. 1172–1176, June 2007.

[13] C. E. Davila and R. Srebro, "Subspace averaging of steady-state visual evoked potentials," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 6, pp. 720–728, 2000.

[14] M. Middendorf, G. McMillan, G. Calhoun, and K. Jones, "Brain-computer interfaces based on the steady-state visual-evoked response," *IEEE Trans. Neural Sys. Rehab. Eng.*, vol. 8, no. 2, pp. 211–214, 2000.

[15] A. Bauer, J. W. Kantelhardt, A. Bunde, P. Barthel, R. Schneider, M. Malik, and G. Schmidt, "Phase-rectified signal averaging detects quasi-periodicities in non-stationary data," *Physica A*, vol. 364, pp. 423–434, 2006.

[16] J. Wolpaw, N. Birbaumer, W. Heetderks, D. McFarland, P. Peckham, G. Schalk, E. Donchin, L. Quatrano, C. Robinson, and T. Vaughan, "Brain-computer interface technology: a review of the first international meeting," *IEEE Trans. Neural Sys. Rehab. Eng.*, vol. 8, no. 2, pp. 164–173, Jun 2000.

[17] B. Allison, D. McFarland, G. Schalk, S. Zheng, M. Jackson, and J. Wolpaw, "Towards an independent brain-computer interface using steady state visual evoked potentials," *Clin Neurophysiol*, vol. 119, no. 2, pp. 399–408, Feb 2008.

[18] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification II edition*. John-Wiley and Sons, 2001.

[19] U. Hoffmann, E. J. Fimbel, and T. Keller, "Brain-computer interface based on high frequency steady-state visual evoked potentials: A feasibility study," in *Proc. 4th Int. IEEE/EMBS Conf. Neural Engineering NER '09*, 2009, pp. 466–469.