Detection of attention shift for asynchronous P300-based BCI

Yichuan Liu, Hasan Ayaz, Adrian Curtin, Patricia A. Shewokis, Banu Onaral

Abstract—**Brain-computer interface (BCI) provides patients suffering from severe neuromuscular disorders an alternative way of interacting with the outside world. The P300-based BCI is among the most popular paradigms in the field and most current versions operate in synchronous mode and assume participant engagement throughout operation. In this study, we demonstrate a new approach for assessment of user engagement through a hybrid classification of ERP and band power features of EEG signals that could allow building asynchronous BCIs. EEG signals from nine electrode locations were recorded from nine participants during controlled engagement conditions when subjects were either engaged with the P3speller task or not attending. Statistical analysis of band power showed that there were significant contrasts of attending only for the delta and beta bands as indicators of features for user attendance classification. A hybrid classifier using ERP scores and band power features yielded the best overall performance of 0.98 in terms of the area under the ROC curve (AUC). Results indicate that band powers can provide additional discriminant information to the ERP for user attention detection and this combined approach can be used to assess user engagement for each stimulus sequence during BCI use.**

I. Introduction

Brain-computer interface (BCI) is an emerging and rapidly growing research area that enables the development of systems that bypass the brain's normal communication pathways of nerve and muscle, allowing the brain to communicate directly with the external world. Clinical applications of BCIs are targeted for patients suffering from severe neuromuscular disorders to provide them with an alternative way of interacting with the world.

The P300-based BCI is among the most popular paradigm in the field due to its ease of use, high performance and reliable signal it can offer [1]. Classification of signals in BCI relies on the P300 event-related potential (ERP) that is elicited in the oddball paradigm. In a P300-based BCI, sequences of visual or audio stimuli are presented to the users of the BCI who are asked to focus their attention on the occurrence of rare target stimuli among more frequent non-targets. The P300 waves are generated by the brain after the user recognized the target stimuli. These P300 waves are then detected and translated to perform actions such as turning on a switch in environmental control, or choosing a letter during a spelling task [1].

Manuscript received March 15, 2012. Y. Liu, H. Ayaz, A. Curtin and B. Onaral are with the Drexel University, School of Biomedical Engineering, Science and Health Systems, Philadelphia, PA 19104 USA. P. Shewokis is with the College of Nursing and Health Professions, Drexel University, Philadelphia, PA 19104 , USA Phone: 215-571-3709; Fax: 215-571-3718; e-mail: (yl565, ayaz, abc48, shewokis, banu.onaral)@drexel.edu.

Research on the P300-based BCI has been mainly focused on three areas: 1) Stimulus presentation paradigm, which concerns designing different presentation modes such as row/column, single cell and checkerboard for faster communication speed, better ERP signal to noise ratio or to enhance user experience. 2) Feature extraction and classification algorithms, which involves developing algorithms to translate EEG signals into actions more accurately, and 3) novel applications. A detailed review of the current status and future directions of P300-based BCI research can be found in [1]. Most BCI systems proposed in the literature are synchronized, requiring that users must follow the pace of the BCI system and that they have no control of when to start and stop using the BCI. A more practical BCI should allow users to interact with it in an asynchronous manner. To achieve this goal of interactive asynchrony, it is critical to determine whether or not the user is attending to the BCI system.

An asynchronous P300-based BCI was first proposed by Zhang et al [2] and studied by various other groups [3-6]. These studies can be grouped into two categories according to the methods applied for user attention detection: those based on statistical analysis of the P300 wave amplitude features, such as [2-4]; and those based on a hybrid BCI, such as in [5] where the steady state visually evoked potentials (SSVEP) paradigm were applied in conjunction with the P300 paradigm, and in [6] where the event-related desynchronization (ERD) paradigm was employed.

In this study, we investigated using band powers as features for user attention detection. It has been long known in the literature that ERP is related to the rhythmic activity of the brain [7-10]. However, most of the studies investigated rhythmic activity using time-frequency analysis during the time course of P300 activity. It is still unknown whether band powers are capable of characterizing whether a user is attending (engaged actively) to the P300-based BCI. To investigate this band power active engagement, we first compared the delta (0.5-4Hz), theta (4-8Hz), alpha (8-13Hz) and beta (13-30Hz) activities of subjects when they were actively engaged in the P300-based BCI to those times when they were not engaged. Secondly, we used the rhythmic activities of selected bands and channels as feature for user attention classification.

II. Materials and Methods

A. Participants

Ten right-handed volunteers (ages between 20 to 24 years) participated in the study but one subject was excluded due to missing synchronization communication. All participants stated that they have no neurological or psychiatric history and gave written informed consent approved by the institutional review board of Drexel University for the experiment. All participants reported that they had no prior experience with BCI.

B. *Experiment setup*

All data acquisition was inside a Faraday cage for electromagnetic and sound insulation. Participants sat comfortably inside the Faraday cage with two monitors positioned on a desk in front of them. The right hand side monitor displayed the customized P300 stimulus which was implemented using the BCI2000 framework [11]. The left hand side monitor displayed the task outcome, which was in this case a 3-D virtual maze rendering from a first person view implemented using MazeSuite software [12]. The P300 stimulus was presented in a 3×3 matrix that contained iconic representation of nine actions: move forward, move backward, turn right, turn left, slide right, slide left, trigger, jump and look around. We described the details of utilizing P300 for spatial navigation control in [13].

C. *Recording*

EEG data were recorded using Neuroscan Synamps 2 from nine locations according to the 10-20 international system: FCz, Cz, CP3, CPz, CP4, P3, Pz, P4, and Oz based on the performance evaluation in [14]. The reference and ground position used are A1 and A2, respectively. Data were sampled at 1000Hz and band pass filtered from 0.5 to 60Hz. A notch filter was also applied at 60Hz to remove possible noise from power source.

D. *Protocol*

Visual stimuli were generated by randomly intensifying columns and rows of the P3speller matrix provided by BCI2000. The stimulus duration was 80ms and the inter-stimulus interval (ISI) was 160ms so the stimulus onset asynchrony (SOA) was 240ms. The ERP window of each *epoch* was set to be 0-1000ms after the onset of a stimulus. There were three types of epochs: *target epoch*, whose onset was target stimulus; *non-target epoch*, whose onset was non-target stimulus; and *control epoch*, during which the subjects were not attending the P3speller. A *sequence* included six epochs. In a sequence, each column and row of the matrix was intensified exactly one time. A *run* was consisting of several sequences. There were two types of runs: *attended run* and *control run*. During the attended runs, subjects were instructed to focus on one icon of the matrix and count the number of times this icon is intensified. After each single attended run, one action was classified and outputted to the maze. In control runs, subjects focused on the maze screen waiting for instruction or configuring the next action they intended to make during maze navigation. A *session* comprised of several attended and control runs.

There were two experimental stages in our protocol: training and testing, each containing several sessions. At the beginning of each stage, subjects performed a few runs under the instruction of the experimenter to get familiarized with the protocol.

1) *Training.*

There were three repeated sessions in this stage. Each session included 12 attended runs and 12 control runs. Each attended run was followed by a control run. All runs were comprised of 10 overlapping sequences (See Figure 1). In attended runs, subjects focused their attention on P300 stimuli for selection of the action they were instructed to perform during the last control run. In control runs, subjects focused their attention on the maze screen for the instruction of the next action. At the end of each attended and control run, a visual notification was shown on the currently attended screen instructing the subject to direct their attention to the opposing screen.

2) *Testing.*

Prior to testing stage, a classifier was trained based on the stepwise linear discriminant algorithm in BCI2000 for online classification. During the testing stage, subjects went through two sessions. In each session, subjects were instructed to navigate within several mazes using the P3 speller from the starting point to the exits. The first session included three randomly ordered 'T' shape and 'double T' shape mazes. The second session included two large mazes. For the large mazes, directions to the exit were shown on the screen to make the task easier. The control run in the testing stage comprised of only six sequences, for the purpose of reducing the total experiment time.

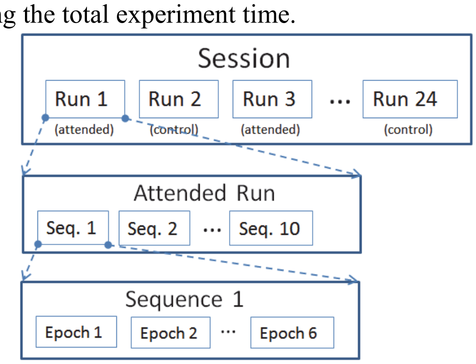

Fig.1 Schematic representation of a session. Each session included 12 attended runs and 12 control runs. Each run included 10 sequences. In each sequence there were 6 flashes (epochs).

E. *Data processing and classification*

The data processing and attendance classification was performed on each single *sequence*. More specifically, it was from the onset of the first stimulus of a sequence to 1000ms after the onset of the sixth (and last) stimulus of the sequence, i.e. $5 \times SOA + 1000 = 2200ms$ data. A threshold based approach (peak-to-peak amplitude) has been applied to reject artifact containing sequences [15]. We employed three methods for attendance classification which was dependent on the different features used for characterization:

1) *Band power*

We estimated power spectral density for each sequence using a periodogram. A two-way repeated measures Analysis of Variance (ANOVA) analysis with Attendance (2 levels: attending, not-attending) and Band (4 levels: delta, theta, alpha and beta) was calculated for each electrode signal. Geisser–Greenhouse correction was used for sphericity and Tukey's post hoc tests to determine the locus of band power main effects. Post-hoc comparisons of interactions were computed using planned contrasts of attendance at each band power. For classification, stepwise linear discriminant analysis (SWLDA) was applied to reduce over-fitting. Other common classification algorithms such as multilayer perceptron, support vector machine and lasso regression have been investigated with similar results. For all tests, the significance criterion set at $\alpha=0.05$.

2) *ERP score.*

Each sequence of data was further divided into six overlapping epochs from the onset of each stimulus. To downsample the signal, we calculated the averages of non-overlapping 50ms data blocks for each epoch. The target epochs and non-target epochs were labeled as 1, the no-control epochs were labeled as 0. The SWLDA was used to train classifiers for discriminating the control epochs from the rest. The aim of this was not to classify the target epochs as in [3] but to detect overall presence of any P300 within the sequence. The ERP score for each sequence was defined as follows:

$$s_{ERP} = \max_{i=1,2,3} s_{row}^{i} + \max_{i=1,2,3} s_{col}^{i}$$

Where s_{row}^{i} and s_{col}^{i} were scores given by the classifiers for the ith row and column epoch, respectively.

3) *Hybrid.*

In the hybrid method, we used the logarithmic band power features selected in 1) and the ERP scores given in 2) for characterization. A SWLDA algorithm was then applied for classification.

III. RESULTS

A. Statistical Comparisons

The band powers from each run were subjected to a two-way repeated measures ANOVA analysis. The results indicated significant main effects of Band (delta, theta, alpha and beta) and Attendance and interaction of Band with Attendance. Significant results are reported in Table 1.

TABLE I
STATISTICAL ANALYSIS

Electrode	Band	Attendance	Interaction	Contrast1	Contrast4
	F (3,24)	F (1,24)	F(3,24)	T (55) *Delta*	T (55) *Beta*
FCz	61.71^	46.64^	27.93^	6.614^	NS
Cz	45.17^	43.30^	21.30^	4.885^	NS
CPz	33.72^	48.03^	13.92#	3.731^	NS
CP3	26.92^	50.17^	35.55^	4.834^	NS
CP4	29.59^	32.27^	13.45#	3.469#	NS
Pz	30.32^	71.90^	12.99#	3.608^	NS
P3	22.95^	71.54^	25.21^	4.116^	NS
P4	32.22^	58.18^	11.37#	3.459#	NS
Oz	31.40^	71.47^	14.63#	4.909^	2.138*

Note: * $p < 0.05$, # $p < 0.01$, ^ $p < 0.001$

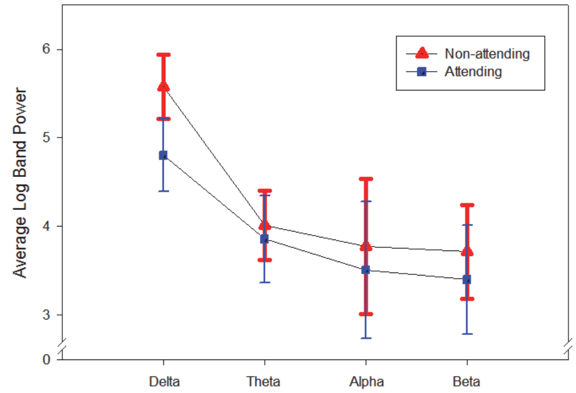

Fig. 1. Average band power from all subjects depicts a significant difference in Delta band. This representative figure from Cz electrode is similar to other electrodes. Error bars are standard deviations.

B. Attendance classification

For comparing the effectiveness of the three classification methods, we first performed a 3-fold cross-validation on the training data (using two data sessions for classifiers training, one data session for making predictions). Then, we trained classifiers on the training data and made predictions on the testing data. We used the area under curve (AUC) of receiver operating characteristic (ROC) curve as the criterion for performance evaluation. In this classification, the true positive represents the rate of attended runs that have been successfully detected. False positive exemplifies the rate of control runs that have been incorrectly classified as attended runs. The results are reported in Tables II and III.

TABLE II
AUC VALUES FROM 3-FOLD CROSS-VALIDATION OF TRAINING DATA

Subject	Band Power	ERP score	Hybrid
1	0.963	0.988	0.996
2	0.903	0.884	0.954
3	0.950	0.915	0.974
4	0.932	0.986	0.992
5	0.969	0.959	0.986
6	0.852	0.939	0.962
7	0.943	0.978	0.990
8	0.937	0.920	0.977
9	0.988	0.973	0.997
Avg.	**0.937**	**0.949**	**0.981**

TABLE III
AUC VALUES FROM TESTING DATA USING CLASSIFIERS TRAINED ON TRAINING DATA

Subject	Band Power	ERP score	Hybrid
1	0.891	0.987	0.992
2	0.950	0.864	0.966
3	0.923	0.911	0.950
4	0.905	0.967	0.979
5	0.987	0.970	0.995
6	0.919	0.959	0.980
7	0.889	0.960	0.974
8	0.978	0.906	0.983
9	0.975	0.879	0.984
Avg.	**0.935**	**0.933**	**0.978**

Results indicate that performance of the band power method was comparable to that of ERP score method whereas the hybrid method provides the best classifications for each subject.

Next, we evaluated the classification threshold for a false positive value smaller than 0.02 based on the ROC curves of the training data cross-validation results and made predictions on the testing data. For the maze navigation task a low false positive is desirable. For other tasks, this constraint can be relaxed to allow a higher true positive rate. The results were shown in Table IV and Table V. Overall, the band power method gave satisfactory false positives but poor true positives. In contrast, the ERP score gave poor false positive but satisfactory true positive. The hybrid method gave the most satisfactory results in terms of both false positive and true positive.

TABLE IV
ATTENDANCE CLASSIFICATION FOR TRUE POSITIVE

Subject	Band Power	ERP score	Hybrid
1	0.21	0.96	0.91
2	0.38	0.46	0.57
3	0.60	0.46	0.66
4	0.38	0.81	0.82
5	0.46	0.80	0.83
6	0.14	0.60	0.71
7	0.28	0.74	0.74
8	0.48	0.50	0.84
9	0.47	0.59	0.81
Avg.	**0.38**	**0.65**	**0.76**

TABLE V
ATTENDANCE CLASSIFICATION FOR FALSE POSITIVE

Subject	Band Power	ERP score	Hybrid
1	0.000	0.071	0.012
2	0.013	0.028	0.003
3	0.021	0.007	0.014
4	0.014	0.014	0.000
5	0.000	0.011	0.004
6	0.004	0.007	0.000
7	0.024	0.018	0.018
8	0.000	0.036	0.000
9	0.004	0.081	0.017
Avg.	**0.009**	**0.031**	**0.007**

IV. DISCUSSION

In this study, we investigated band powers as potential features to detect user's attendance toward a P3speller. Results showed that user attendance characterization performance of band powers were comparable to the conventional approach (that is by using P300 wave amplitudes). A hybrid classification using both features yielded the best overall performance (See Table II and III) which suggests that there is complementary independent information in each of these features. It has been shown that the evoked P300 waveform in visual/auditory oddball paradigms has a frequency characteristic in the delta–theta range [7, 16]. Most studies focused on single trial analysis, however, in this study, our aim was to look at a larger window and utilize frequency characteristics that evolve over a *sequence* instead of a single trial (See Fig 1) with the most important difference due to attendance change in the delta band.

The user attendance detection approach proposed in this study can be implemented to run in parallel with a P3 speller target detection algorithm to improve the BCI reliability and usability. The scores given by the new classifier may serve as an index of how actively the user is engaged with the P300 speller. Moreover, this approach can be used to build self-paced / asynchronous BCIs. However, further investigation is required to include control runs when subjects are engaged in different tasks such as arithmetic, listening to music and so forth to test the robustness of the proposed algorithm with other types of BCI. Future work will include testing the proposed algorithm online and also with clinical subject populations such as ALS patients and in real world environments.

REFERENCES

[1] J. Mak, Y. Arbel, J. Minett *et al.*, "Optimizing the P300-based brain–computer interface: current status, limitations and future directions," *Journal of Neural Engineering,* vol. 8, pp. 025003, 2011.

[2] H. Zhang, C. Guan, and C. Wang, "Asynchronous P300-Based Brain--Computer Interfaces: A Computational Approach With Statistical Models," *Biomedical Engineering, IEEE Transactions on,* vol. 55, no. 6, pp. 1754-1763, 2008.

[3] F. Aloise, F. Schettini, P. Aricò *et al.*, "P300-based brain–computer interface for environmental control: an asynchronous approach," *Journal of Neural Engineering,* vol. 8, pp. 025025, 2011.

[4] A. Lenhardt, and H. Ritter, "An augmented-reality based brain-computer interface for robot control," *Neural Information Processing. Models and Applications*, pp. 58-65, 2010.

[5] R. Panicker, S. Puthusserypady, and Y. Sun, "An asynchronous p300 bci with ssvep-based control state detection," *Biomedical Engineering, IEEE Transactions on*, no. 99, pp. 1-1, 2011.

[6] H. Riechmann, N. Hachmeister, H. Ritter *et al.*, "Asynchronous, parallel on-line classification of P300 and ERD for an efficient hybrid BCI." pp. 412-415.

[7] C. Basar-Eroglu, E. Basar, T. Demiralp *et al.*, "P300-response: possible psychophysiological correlates in delta and theta frequency channels. A review," *International journal of psychophysiology,* vol. 13, no. 2, pp. 161-179, 1992.

[8] C. Basar-Eroglu, and T. Demiralp, "Event-related theta oscillations: an integrative and comparative approach in the human and animal brain," *International journal of psychophysiology,* vol. 39, no. 2-3, pp. 167-195, 2001.

[9] E. M. Bernat, S. M. Malone, W. J. Williams *et al.*, "Decomposing delta, theta, and alpha time-frequency ERP activity from a visual oddball task using PCA," *International journal of psychophysiology,* vol. 64, no. 1, pp. 62-74, 2007.

[10] X. Wang, and M. Ding, "Relation between P300 and event-related theta-band synchronization: A single-trial analysis," *Clinical Neurophysiology,* vol. 122, no. 5, pp. 916-924, 2011.

[11] G. Schalk, D. J. McFarland, T. Hinterberger *et al.*, "BCI2000: a general-purpose brain-computer interface (BCI) system," *Biomedical Engineering, IEEE Transactions on,* vol. 51, no. 6, pp. 1034-1043, 2004.

[12] H. Ayaz, P. A. Shewokis, A. Curtin *et al.*, "Using MazeSuite and functional near infrared spectroscopy to study learning in spatial navigation," *J. Vis. Exp., DOI,* vol. 10, pp. 3443, 2011.

[13] A. Curtin, H. Ayaz, Y. Liu *et al.*, "A P300-based EEG-BCI for Spatial Navigation Control," in Conf Proc IEEE Eng Med Biol Soc, (in press), 2012.

[14] D. Krusienski, E. Sellers, D. McFarland *et al.*, "Toward enhanced P300 speller performance," *Journal of neuroscience methods,* vol. 167, no. 1, pp. 15-21, 2008.

[15] R. J. Croft, and R. Barry, "Removal of ocular artifact from the EEG: a review," *Neurophysiologie Clinique/Clinical Neurophysiology,* vol. 30, no. 1, pp. 5-19, 2000.

[16] V. Kolev, T. Demiralp, J. Yordanova *et al.*, "Time-frequency analysis reveals multiple functional components during oddball P300," *Neuroreport,* vol. 8, no. 8, pp. 2061, 1997.