An Exploratory Study of Factors Affecting Single Trial P300 Detection

Ben H. Jansen*, *Senior Member, IEEE*, Anand Allam, Prashant Kota, Kathleen Lachance, Ayokunle Osho, and Karthik Sundaresan

***Abstract*—A threshold detector for single-trial P300 detection has been evaluated. The detector operates on the 0–4 Hz band, isolated from the raw electroencephalogram using low-pass filtering, wavelet transforms, or the piecewise prony method (PPM). A detection rate around 70% was found, irregardless of stimulus type, interstimulus interval (ISI), probability of occurrence (P_r) of the target stimuli, intrasession and intersession effects, or filtering method. This suggests that P300-based brain-machine interfaces can use an ISI as short as 1 s and a P_r of 45%, to increase throughput.**

***Index Terms*—Brain-machine interface, evoked responses, single-trial analysis, P300.**

I. Introduction

THE P300 is one type of brain activity that can be used for a brain machine interface (BMI). The P300 is a positive component occurring about 300 ms poststimulus [1] and is typically elicited using an "oddball" paradigm. This involves presenting a sequence of two types of sensory stimuli, with one (the "target") occurring less frequently than the other. If the subject is required to attend to the target stimuli (e.g., by counting them), a P300 will be visible in the ensemble average of the responses to the target stimuli, but not in the response to the nontargets. The P300 is maximal over the parietal cortex [2], [3], and is thought to reflect the mental operations preceding response selection and execution [4].

The feasibility of using the P300 for a brain-machine interface (BMI) was demonstrated by Donchin and co-workers [5], [6], who developed a spelling device. Unfortunately, the throughput of a P300-based BMI may be as low as 1 character/min. This is partially due to the fact that ensemble averaging ten or more trials (as done in [5]) is required to improve the signal-to-noise ratio because the evoked activity is buried in the ongoing, spontaneous electroencephalogram (EEG). Hence, the use of single-trial methods could substantially increase the communication speed. Other factors affecting the bit rate are the interstimulus interval (ISI), which is typically set around 2 s, and the probability of occurrence (P_r) of the rare stimuli (typically around 15%). Shorter ISIs and larger P_r could lead to a further increase in throughput. Another aspect that needs study is how the detection rate is affected by short-term changes (habituation) and long-term changes (learning) in P300 morphology that may occur.

The effect of ISI and P_r on P300 generation has been studied before, but those studies focused primarily on changes in amplitude and latency, and not on detection rates. For example, reduced P300 amplitudes and increased latencies have been reported for 0.5 s ISIs [7]. Polich [8] showed that the P300 amplitude decreases with an increase in target stimulus probability for ISIs ranging from 2 to 3 s but exhibited no difference for ISIs of 4 to 10 s. Donchin and Coles [9] found that P300 amplitude decreases with an increase in target-stimulus probability for a wide range of target probability manipulations.

Several studies suggest that the P300 amplitude diminishes over the course of an experiment (habituation) [10]–[13]. Few studies of the long-term patterns of change in P300 have been performed. Kinoshita *et al.* [14] report a significant gradual decrease in P300 amplitude over five sessions at intervals of 7 to 10 days, followed by a partial recovery one week later. Recordings made one month later showed an increase in P300 amplitude, possibly indicative of a learning effect.

In this paper, we study how the performance of a single-trial P300 detection method is affected by the choice of ISI and P_r, and how its detection rate fluctuates within and between sessions. Such knowledge is important for the development of P300-based BMIs. The detector consists of a decomposition stage that separates the delta band (0–4 Hz) from the raw EEG, followed by a threshold detector. Better than 70% correct detection rates have been reported for such a detector using low-pass filtering [15], or wavelet transforms [16]. We will compare these decomposition methods to the piecewise prony method (PPM) which was developed by us previously for single-trial EP analysis [17].

II. Methods

Sixteen channels (Fp1, Fp2, F3, Fz, F4, T7, T8, C3, Cz, C4, P3, Pz, P4, O3, Oz, and O4) of EEG were recorded from 4 normal, young adults (3 male, 1 female) between the ages of 20 and 25 years using the ActiveTwo system (BioSemi, The Netherlands), which includes active electrodes attached to the scalp using an elastic head cap. The average of all the channels

Manuscript received August 13, 2003; revised January 16, 2004. This work was supported in part by the National Institute of Mental Health under Grant 1 R01 MH58784, in part by a GEAR grant from the University of Houston, and in part by the National Science Foundation (NSF) under Grant 0240889. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF, NIH, or UH. *Asterisk indicates corresponding author.*

*B. H. Jansen is with the Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77204-4005 USA (e-mail: bjansen@uh.edu).

A. Allam is with the Department of Electrical Engineering, University of Pennsylvania, Philadelphia, PA 19104-6391 USA.

P. Kota, A. Osho, and K. Sundaresan are with the Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77204-4005 USA.

K. Lachance is with the Department of Engineering Science, Trinity University, San Antonio, TX 78212-7000 USA.

Digital Object Identifier 10.1109/TBME.2004.826684

served as the reference. The EEG was sampled at 2048 Hz, subjected to baseline correction to remove any dc drift, and filtered between 0.1 Hz and 40 Hz. The data were downsampled to 250 Hz before further processing took place.

Two variants of the visual oddball paradigm were used, with one involving the display of a sequence of "X" and "O" characters with the "X" being the target, while right and left arrow symbols were used in the second paradigm with the right arrow being the target. The left/right arrow paradigm was used for subject 1 exclusively, while the other three subjects underwent the X/O paradigm. Each experiment consisted of 200 or 400 stimuli, with an ISI of 1 s or 2 s, and a P_r of 15% or 45%. Data were obtained over periods spanning three weeks to three months. Subjects were instructed to respond to target stimuli by clicking the left button of the mouse. Trials in which the amplitude exceeded 75 μV were discarded because of possible artifacts. There remained responses to at least 10 target stimuli in each session (maximum 87, median 48, mean 45), and at least 14 nontargets (maximum 323, median 119, mean 153).

The detector operates on the 0–4 Hz frequency band (delta activity) obtained from the Pz channel, since P300 is primarily associated with activity in that band [18] and maximal over the parietal cortex when evoked using visual stimuli [2], [3]. The delta activity was separated from the EEG using a low-pass digital filter (LP), wavelet transformation (WT), or the PPM. A ninth-order Chebyshev type II (ripple in the stopband) filter with a pass band frequency of 3.5 Hz, a stop band frequency of 4 Hz, and a stop band ripple of 25 dB was used to isolate the 0–4 Hz band. The EEG data were processed twice with this filter: the output of the first filtering step was time-inverted and filtered again to obtain zero-phase shift. The wavelet transform used a dyadic decomposition scheme to obtain the 0–4 Hz activity. The Daubechies-5 wavelet was used because the asymmetrical shape of this wavelet has some resemblance to the P300 component. The piecewise prony method (PPM) [17] models the activity in a specific frequency band by means of exponentially increasing and/or decreasing sinusoidal components ("Prony components"). The amplitude, frequency, phase, start and end point of each of these components is determined by minimizing a mean-squared error criterion.

The three decomposition methods were applied to each of the single trials, starting 0.5 s pre stimulus through 1.5 s poststimulus, and the outputs subjected to a threshold detector. The threshold was determined for each <Asession separately using the first 20 nontarget trials in a session. The value of the threshold was set to the 95 percentile of the prestimulus amplitude values (i.e., the output of the decomposition stage) derived from these trials. A P300 was indicated if the following two criteria were met. First, the threshold had to be exceeded between 350 and 500 ms poststimulus. This time interval was selected on the basis of the P300 latency observed in our data. Second, the local minimum between the 0 ms mark and the location of the peak must be below the threshold. This ensures that a signal with a high dc and no P300 is not mistaken for a target response.

III. Results

The PPM-based threshold detector correctly identified 70% of all the 1580 target responses, and misclassified 30% of the 5 353 nontarget responses for an overall index of merit of 70% over all 35 experiments (minimum 45%, maximum 87%, median 70%). The low-pass filter-based approach produced the same index of merit, and the wavelet transform resulted in an index of merit of 69%. These differences were not significant (paired t-test, $p = 0.98$ for PPM versus LP, $p = 0.23$ for WT versus LP, and $p = 0.27$ for PPM versus WT).

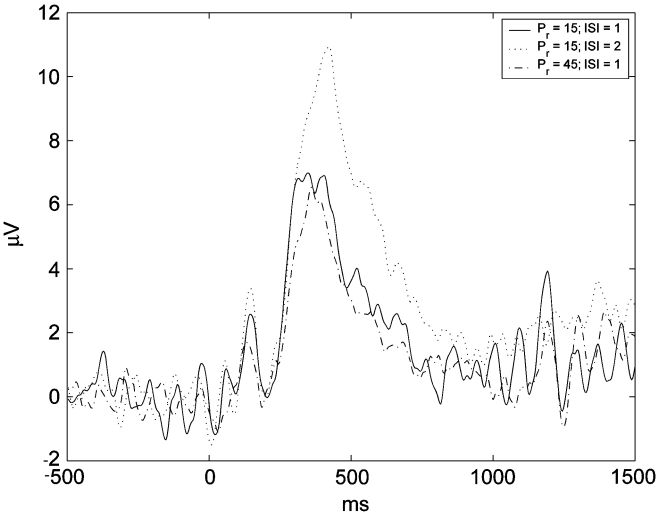


Fig. 1. Grand ensemble averages for the rare trials for P_r of 15% and 45% (both with $\mathrm{ISI} = 1$ s), and $P_r = 15\%$ with $\mathrm{ISI} = 2$ s.

The effect of varying P_r was studied by means of a pair-wise comparison (t-test) between the index of merit obtained for a P_r of 15% and a P_r of 45% (ISI = 1 s for both). The PPM was used to extract the delta activity. Each pair of recordings was obtained on the same day, and we had eight such pairs. The index of merit was 67% $(\sigma = 6)$ for $15\% P_r$, and 69% $(\sigma = 9)$ for 45% P_r. The difference between the indexes of merit were not significant $(t = 0.7, p = 0.51)$. The grand ensemble averages for the two conditions, presented in Fig. 1, are very similar, which explains the similarity between the indexes of merit.

The ISI effect was studied in a similar manner, using pairs of data, obtained on the same day, with a 2 s and 1 s ISI, respectively, and $P_r = 15\%$. Again, eight such pairs were available. The 2 s ISI data produced an index of merit of 74% $(\sigma = 5)$, and 67%$(\sigma = 6)$ was obtained for a 1 s ISI. This difference is almost significant at the 95% level $(t = 2.3, p = 0.055)$. The grand average for the 2 s ISI shows a substantially larger P300 than for the 1 s ISI, as one can see from Fig. 1.

Habituation and long-term effects were studied using the data from two of the four subjects (subject 1 and 2), obtained for 2 s ISI and 15% P_r. Five data sets were available for each subject, spanning a 3-mo and 2-mo period, respectively. The PPM-based detector was applied and the index of merit was obtained for each session. Furthermore, a separate index of merit was computed for the first 20 target responses and the nontargets occurring over that time span, and the second 20 target responses (and corresponding nontargets). All three indexes were computed using the same threshold obtained from the first 20 nontargets of a session. A paired t-test did not find significant differences between the indexes of merit computed for the first and the second "half" of the sessions $(t = 0.236, p = 0.472)$. The

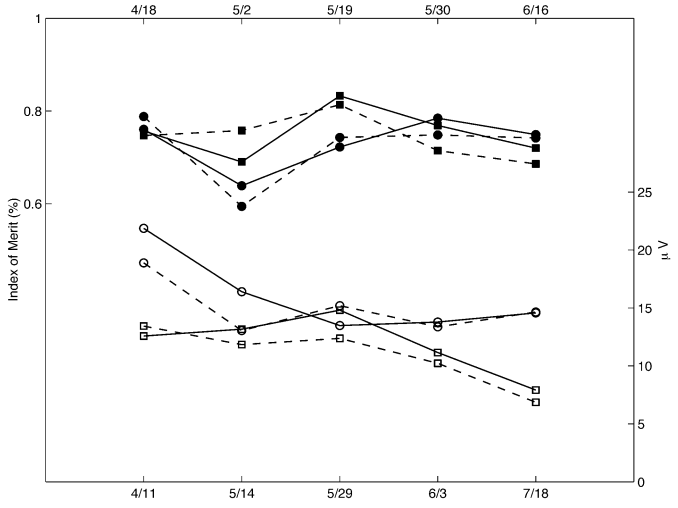


Fig. 2. Index of merit (left scale, solid squares and circles) and P300 amplitude (right scale, open squares and circles) as a function of time. Results for subject 1 are indicated by circles, and the recording dates are shown along the bottom scale. Results for subject 2 are indicated by squares, with the recording dates shown along the top axis. Indexes of merit are shown for the first 20 target stimuli and corresponding nontargets (indicated by solid lines), and for the second 20 target and corresponding nontargets (indicated by dashed lines). P300 amplitudes for the first 20 target stimuli are indicated by solid lines, and dashed lines are used for the second 20 targets.

F-test indicated that the inntrasession variability was smaller than the intersession variability ($p = 2.17 \cdot 10^{-6}$), but there was no hint of a trend in the data suggesting that later sessions produced higher (or lower) detection rates than earlier sessions, as can be seen from Fig. 2. However, the P300 amplitudes (measured from peak to immediately preceding trough) were generally lower for the second 20 target responses than the first 20 (12.99 μV versus 13.98 μV), but this difference was not significant ($t = 1.92, p = 0.087$). Both subjects showed a considerable decrease in P300 amplitude as time progressed, as can be seen from Fig. 2. A regression analysis did not find a significant correlation between P300 amplitude and P300 detection rate ($p = 0.816$) or index of merit ($p = 0.38$).

The effect of stimulus type (i.e., arrow or letter) was evaluated using a t-test for differences between two means. The data for subject 1 and 2 were used with $P_r = 15\%$ and ISI of 2 s. Subject 2 had a slightly larger index of merit (74%, $\sigma = 4$) than subject 1 (70%, $\sigma = 7$), suggesting that the "X" target produced larger responses than the right arrow, but the difference was not significant ($t = -1.13, p = 0.33$).

Our subjects were required to respond to target stimuli by a mouse click, and one may argue that the associated motor activity facilitates P300 detection. This was tested using three pairs of data sets from subject 1, obtained over three different days, with each pair consisting of one recording using the standard paradigm (i.e., mouse click), and a paradigm where the subject was told to mentally count the occurrence of the target stimuli. An ISI of 2 s, and a P_r of 15% was used. The difference between the correct detection rate for the paradigm requiring a mouse click ($82\% \pm 0.1$) and no mouse click ($78\% \pm 0.08$) was not significant ($t = 0.78, p = 0.26$).

We also explored to what degree discriminating the same "event" from one trial to another is independent using one data set from each of the four subjects (ISI = 1 s, $P_r = 45\%$). Single trials were classified as before, but the decision whether a trial was target or nontarget was made only when at least two of three successive target responses (or nontarget responses) were classified as target (nontarget). This led to no change in target detection rate, but a significant decrease in false detection rate (from 23% to 13%, $t = 2.44, p < 0.046$).

IV. Discussion

Our findings suggest that probability of occurrence has little or no effect on the ability to differentiate between responses to target and nontarget stimuli. P300 amplitude was the same for both conditions, contradicting a previous report of an inverse relationship between P300 amplitude and P_r [9]. The 2 s ISI produced marginally larger detection rates, but the increase in performance over the 1 s ISI was not significant, even though the P300 amplitude was substantially larger for the 2 s ISI. The larger P300 amplitude observed for the 2 s ISI matches previous reports [7]. Neither short-term "habituation" nor long-term "learning" effects were found in the index of merit, although substantial P300 amplitude differences (not unlike those reported in [14]) were seen. This indicates that the threshold used for P300 detection is well below the P300 amplitude. We did not find a significant difference in detection rate between the paradigm requiring a motor response to target stimuli and one that required mental counting instead, suggesting that the motor response does not significantly contribute to P300 detection. Our results also suggest that there is some degree of independence between successive responses, especially in case of the nontarget stimuli.

Overall, the index of merit hovers around 70%, irregardless of stimulus parameters, ISI, P_r, or decomposition method. A similar performance has been reported in [15], [16]. However, one should not draw the conclusion that 70% is the best one can do with regard to P300 detection. After all, the focus of this study was on stimulus-related factors that may affect P300 detection, and not on developing the "optimal" single-trial P300 detector. We are currently working on more elaborate detectors, using multiple features, including ones related to the latency and morphology of the P300.

In conclusion, our results suggest that single-trial P300 detection is hardly affected by an ISI as short as 1 s or a P_r as large as 45%. These results should be helpful in increasing the throughput of a P300-based BMI.

References

[1] R. Spehlmann, *Evoked Potential Primer*. Stoneham, MA: Butterworth, 1985.

[2] S. Sutton, M. Braren, J. Zubin, and E. R. John, "Evoked-potential correlates of stimulus uncertainty," *Science*, vol. 150, pp. 1187–1188, 1965.

[3] W. G. Walter, R. Cooper, V. J. Aldridge, W. C. McCallum, and A. L. Winter, "Contingent negative variation: An electric sign of sensorimotor association and expectancy in the human brain," *Nature*, vol. 203, pp. 380–384, 1964.

[4] G. McCarthy and E. Donchin, "A metric for thought: A comparison of P300 latency and reaction time," *Science*, vol. 211, pp. 77–80, 1981.

[5] E. Donchin, K. M. Spencer, and R. Wijesinghe, "The mental prosthesis: Assessing the speed of a P300-based-brain-computer interface," *IEEE Trans. Rehab. Eng.*, vol. 8, pp. 174–179, June 2000.

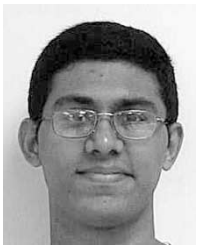
[6] L. A. Farwell and E. Donchin, "Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials," *Electroenceph. Clin. Neurophysiol.*, vol. 70, pp. 510–523, 1988.
[7] D. L. Woods and E. Courchesne, "The recovery functions of auditory event-related potentials during split-second discrimination," *Electroencephalogr. Clin. Neuophysiol.*, vol. 65, pp. 304–315, 1986.
[8] J. Polich, "P300, probability and interstimulus interval," *Psychophysiology*, vol. 27, pp. 396–403, 1990.
[9] E. Donchin and M. G. H. Coles, "Is the P300 component a manifestation of context updating?," *Behav. Brain Sci.*, vol. 11, pp. 357–374, 1988.
[10] G. S. Lew and J. Polich, "P300, habituation, and response mode," *Physiol. Behav.*, vol. 53, pp. 111–117, 1993.
[11] D. Ravden and J. Polich, "Habituation of P300 from visual stimuli," *Int. J. Psychophysiol.*, vol. 30, pp. 359–365, 1998.
[12] D. Ravden and J. Polich, "On P300 measurement stability: Habituation, intra-trial block variation, and ultradian rhythms," *Biol. Psychol.*, vol. 51, pp. 59–76, 1999.
[13] R. Romero and J. Polich, "P(300) habituation from auditory and visual stimuli," *Physiol. Behav.*, vol. 59, pp. 517–522, 1996.
[14] S. Kinoshita, M. Inoue, H. Maeda, J. Nakamura, and K. Morita, "Long-term patterns of change in ERP's across repeated measurements," *Physiol. Behav.*, vol. 60, pp. 1087–1092, 1996.
[15] J. Röschke, K. Mann, M. Wagner, M. Grözinger, J. Fell, and C. Frank, "An approach to single trial analysis of event-related potentials based on signal detection theory," *Int. J. Psychophysiol.*, vol. 22, pp. 155–162, 1996.
[16] T. Demiralp, A. Ademoglu, M. Schürmann, C. Başar-Eroglu, and E. Başar, "Detection of P300 waves in single trials by the wavelet transform," *Brain. Lang.*, vol. 66, pp. 108–128, 1999.
[17] V. Garoosi and B. H. Jansen, "Development and evaluation of the piecewise prony method for evoked potential analysis," *IEEE Trans. Biomed. Eng.*, vol. 47, pp. 1549–1554, Dec. 2000.
[18] B. H. Jansen, A. Hegde, R. Jacob, H. Doshi, and N. N. Boutros, "Evoked potential generation revealed by nonstationary EEG analysis," presented at the *2nd Joint EMBS-BMES Conf.*, Houston, TX, Oct., 23–26 2002.

Ben H. Jansen (S'74–M'76–SM'90) received the B.t.w. and Ir. degrees in electrical engineering from the Technische Hogeschool Twente, Enschede, The Netherlands, and the Ph.D. degree in medical informatics from the Vrije Universiteit, Amsterdam, The Netherlands.

After a two-and-a-half-year stay at Vanderbilt University, Nashville, TN, as a Research Associate and Research Assistant Professor in the Department of Electrical and Biomedical Engineering, he joined the University of Houston, where he is a Professor in the Department of Electrical and Computer Engineering, and a founding member of the Center for Neuro-Engineering and Cognitive Science. His research interests remain focused on EEG and EP analysis.

Anand Allam is a candidate for the B.S.E. degree in electrical engineering and B.A. degree in biology from the University of Pennsylvania, Philadelphia.

He spent the summer at the University of Houston, Houston, TX, in the 2003 NSF REU program working on EEG analysis. He also has experience studying myopia at the Children's Hospital of Philadelphia. After graduation, he plans to study medicine and continue biomedical research.



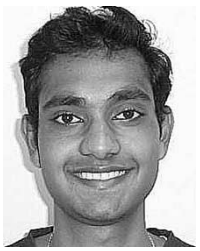
Prashant Kota was born in Chennai, India, and received the Bachelor of Engineering degree from the Madurai Kamaraj University, Madurai, India. He is currently working towards the Masters degree in electrical and computer engineering at the University of Houston, Houston, TX.

Kathleen Lachance is originally from Colorado. She is an undergraduate student at Trinity University, San Antonio, TX. In May 2004, she will receive the B.S. in engineering science with a concentration in electrical engineering.

Upon graduation, she intends to pursue the Ph.D. degree in biomedical engineering.

Ayokunle Osho was born in Lagos, Nigeria, in 1981. He is working towards the B.Sc. degree in electrical and computer engineering at the University of Houston, Houston, TX. During the summer of 2003, he participated in the 2003 NSF REU program working on evoked potential analysis.

Upon graduation, he will seek an advanced degree in electrical engineering.



Karthik Sundaresan was born in Chennai, India. He received the Bachelor of Engineering degree in electronics and telecommunications from the University of Mumbai, Mumbai, India. He is currently pursuing the Masters degree in electrical and computer engineering at the University of Houston, Houston, TX.