Proceedings of the 29th Annual International
Conference of the IEEE EMBS
Cité Internationale, Lyon, France
August 23-26, 2007.

FrP1D6.3

Human Error in P300 Speller Paradigm for Brain-Computer Interface

Reza Fazel-Rezai, *Senior Member, IEEE*

Abstract— **A brain-computer interface (BCI) is a system that conveys messages and commands directly from the human brain to a computer. The BCI system described in this work is based on P300 speller BCI paradigm designed by Farwell and Donchin in 1988. It has been the most widely used and a benchmark in P300 BCI. In this paradigm, a 6x6 matrix of letters and numbers is displayed and subject focuses on a character while different rows and columns flash. The work presented in this paper is an attempt to improve the accuracy of P300 BCI systems by understanding a source of error in this paradigm. It is shown that adjacent rows and columns to the target ones play major role in the error. This can be attributed to human error that when the adjacent row or column to the target one flashes, it attracts subject's attention and creates the P300.**

I. Introduction

Communication and the ability to interact with the environment are basic human needs. Millions of people worldwide suffer such severe physical disabilities that they cannot even meet these basic needs. Even though they may have no motor mobility, however, the sensory and cognitive functions of the physically disabled are usually intact. This makes them good candidates for Brain-Computer Interface (BCI) technology which provides a direct electronic interface to convey messages and commands directly from the human brain to a computer [1]. BCI system involves monitoring conscious brain electrical activity, via electroencephalogram (EEG) signals, and detecting characteristics of EEG patterns, via digital signal processing algorithms, that the user generates to communicate. By establishing a communication link between a subject and a computer, it has the potential to enable the physically disabled to perform many activities, thus improving their quality of life and productivity, allowing them more independence, and reducing social costs. The challenge with BCI, however, is to extract the relevant patterns from the EEG signals produced by the brain each second. EEG signals are voltage changes of tens of microvolts at frequencies ranging from below 1 Hz to about 50 Hz. The inputs to a BCI system are EEG signals recorded from the scalp or brain surface using a specific system of electrode placement called the International 10-20 system developed by Jasper [2].

Different types of brain activity are reflected in EEG signals and have been used in BCI. One of them is the P300 component of Event-Related Potentials (ERPs) [3][4][5]. An ERP is important brain activity that is used in BCI. It is stereotyped electrophysiological response to an internal or external stimulus or simply any measured brain response that is directly the result of a thought or awareness. It is a voltage fluctuation in the EEG induced within the brain that is time-locked to a sensory, motor, or cognitive event. ERPs provide a different indication of how the stimulus is processed. ERPs are measures to reflect the responses of the brain to events in the external or internal environment of the organism. Although they are not easy to be detected, they have wide usage for clinical-diagnostic and research purposes. The ERP response to a stimulus has different components as shown in Fig. 1. It should be noted that whereas the ongoing EEG activity is in the range of 50 microvolts, the ERP is much smaller in the range of 5 to 10 microvolts. The "peaks" and "valleys" in the ERP are termed components, with positive deflections (upward in this picture) labelled with "P" and negative deflections (downward in this picture) labelled with "N". The N400, for example, is the negative peak near 400 milliseconds after stimulus onset. The P300 component is the fourth major positive peak.

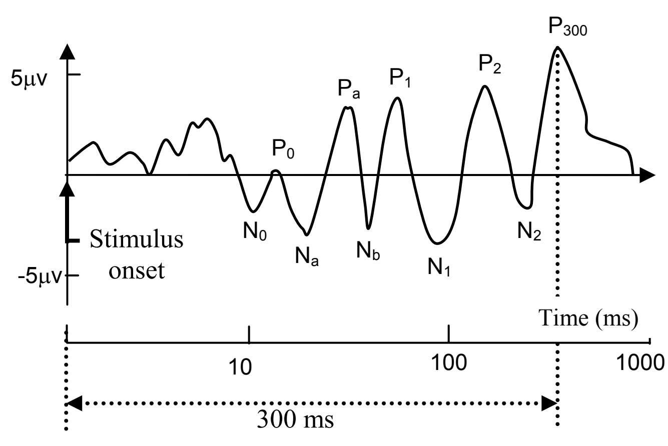

Fig. 1. Components of Event Related Potentials (ERP) after stimulus at time zero [6]. The P300 is the largest component that occurs approximately 300 ms after stimulus onset.

The P300 component of ERP is also the largest component. It is a peak at about 300ms after stimulus, as shown in Fig 1. It is elicited with a simple, two-stimulus discrimination task. This procedure has been dubbed the

Manuscript received April 1, 2007. This work was supported in part by the Department of Electrical and Computer Engineering, the University of Manitoba, Winnipeg, MB R3T 5V6, Canada.

R. Fazel-Rezai is with the Department of Electrical and Computer Engineering, the University of Manitoba, Canada and Institute of Biodiagnostics, National Research Council (IBD-NRC) of Canada (phone: 204-474-9490; fax: 204-261-4639; e-mail: fazel@ee.umanitoba.ca).

"oddball paradigm", whereby two stimuli are presented in a random series such that one of them (the "oddball") occurs relatively infrequently. A user is asked to distinguish between the stimuli by mentally noting each oddball stimulus, which creates a P300 wave, and ignoring the standard stimulus. There are also alternative tasks which can be used to elicit the P300. One of them is the "single stimulus paradigm". In this task, a target stimulus occurs randomly in time, but it is sometimes replaced by silence and the subject is required to respond to every stimulus.

Virtually, any sensory modality such as auditory, visual, somato-sensory, olfactory or even taste stimulation can be used to elicit the response [7]. The shape and latency of the P300 differs with each modality.

Different phenomena can affect the amplitude and latency of the P300 and therefore the performance of this BCI system. In this paper, a source of error in P300 that has not been fully addressed in literature is described. It is shown that adjacent rows and columns are the major source of error in target P300 generation. Experimental evidence for this phenomenon is also presented. This paper is organized as follows. In the next section, methods including the P300 BCI paradigm, data sets and feature extraction and classification are described. In Section 3, results are presented and discussed. Finally, in the last section, concluding remarks are made.

II. Method and Materials

A. *P300 BCI Paradigm*

Farwell and Donchin [8] designed a BCI speller that is based on visual oddball paradigm as shown in Fig. 2. Their speller is the most widely used P300-based BCI since its original design in 1988. In this speller, a 6x6 matrix of symbols, comprising all 26 letters of the alphabet and 10 digits (0-9), is presented to the user on a computer screen randomly at the high speed. At any given moment, the user focuses on the character he/she wishes to communicate, and mentally counts the number of times the selected character flashes. In response to the counting of this oddball stimulus, the row and column of the selected character elicit a P300 wave, while the other 10 rows and columns do not. Detection of the P300 makes it possible to match the responses to one of the rows and one of the columns, and thus identify the target character.

A	B	C	D	E	F
G	H	I	J	K	L
M	N	O	P	Q	R
S	T	U	V	W	X
Y	Z	1	2	3	4
5	6	7	8	9	0

Fig. 2. Farewell and Donchin speller paradigm based on the P300 [8]

This paradigm was an enormous step toward P300 based BCIs and has been a benchmark for P300 BCI systems. One of the greatest advantages of the P300-based BCI is that it does not require intensive user training, as P300 is one of the brain's "built-in" functions. However, P300 detection for real-time applications is not easy. There are several issues to be addressed before any P300-based BCI can be taken outside the research laboratory and put to practical application. One of the important issues is that EEG signal patterns change due to factors such as motivation, frustration, level of attention, fatigue, mental state, learning, and other nonstationarities that exist in the brain. In addition, different users might provide different EEG patterns. All these factors create a need for advanced digital signal processing algorithms to detect the P300 accurately and fast. In spite of all advance signal processing algorithms applied to this paradigm, there has been limited use of the paradigm in real world applications. One of the obstacles has been low accuracy in P300 detection in real time. Several human perceptual phenomena such as attentional blink, repetition blindness, target probability and habituation are potential sources of error in P300 detection and have been addressed in literature.

Raymond *et al.*[9] showed that attentional blink occurs if the interval between two targets is less than 500 msec. In such a case, the first target is correctly identified, while the second is not detected at all. This can be a source of P300 speller error if a non-target row/column near the target attracts attention by flashing less than 500msec before the target is flashed. Attentional blink can be a source of error in this paradigm. Kanwisher [10] showed the effect of repetition blindness. He showed that if two identical targets in a stream of non-targets are flashed at intervals of less than 500ms, the second target may be missed. This can be a source of error whenever a target row (column) is flashed less than 500ms after flashing a target column (row). Although they are intensified in random sequences, because the total intensification plus blank time is 175 msec, there are several cases where the interval between two targets is less than 500ms. This could be another source of error. Donchin *et al.* [11] showed that target probability changes the P300 amplitude. It has been shown that the P300 amplitude is related to the probability of oddball occurrence (target row/column flashing). The less probable the oddball event, the larger the P300 amplitude. Although, in this paradigm, the probability of creating the P300 wave is 0.17 (2/12), the P300 amplitude can be increased by decreasing the oddball probability. Ravden *et al.* [12] showed that attention decreases with repeated presentation of the same stimulus. This phenomenon, habituation, happens when the user loses focus on the target character, the P300 is not elicited, and the system error increases with time.

The purpose of the study described here was to provide empirical evidence for another possible human perceptual error in this paradigm, that is, the effect of adjacent rows and columns in generating P300.

B. Data Sets

In this work, we used the dataset from BCI 2003 competition provided by Blankertz et al. [13], Wadsworth Center, Albany, NY. This dataset was selected, because the results can be compared with the results of other works and can be repeated by other research groups. EEG signals were recorded from 64 electrodes; however, we used only Fz, Pz, Cz, C1, and C2 channels. Row/column intensifications (100msec) were block randomized in blocks of 12. Sets of 12 intensifications were repeated 15 times for each character. Each sequence of 15 sets of intensifications was followed by a 2.5 s period, and during this time the matrix was blank. This period informed the user that this character was completed and to focus on the next character in the word that was displayed on the top of the screen. In other words, for each character, 180 entries of feature values are stored 90 for row intensification and 90 for column intensification. For example, "A" is recognized only when row 7 and column 1 features indicate a P300 component. The signals were digitized at 240Hz and collected from one subject in two sessions. Each session consisted of a number of runs. In each run, the subject focused attention on a series of characters. Target words presented to the subject were: BOWL, CAT, DOG, FISH, FISH, GLOVE, HAT, HAT, RAT, SHOES, and WATER. Note that words FISH and HAT were presented two times. The total number of target characters in these words is 42.

C. P300 Feature Extraction and Classification

Features were extracted from averaged Mexican hat wavelet coefficients. More details of the feature extraction can be found in [14][15][16]. It should be noted that since the objective of this work was to reveal the possible errors, no learning set was used.

III. Results and Discussion

The rare target row and column should elicit a large P300. Target character is associated with cell at the intersection of the target row and column. For a given target character, it is expected to have the P300 in EEG signals related with only one column and one row. The EEG epochs associated with the entire target and non-target stimuli were averaged over all trials and Fz, Pz, Cz, C1, and C2 channels. Two features were calculated for these signals [14][15][16]. From the features, one row and one column were classified as the "detected" row and column. The first and second rows in Table 1 show the target words and characters, respectively. The detected character is shown in the third row. Differences between detected columns/rows and target columns/rows are shown in the fourth and fifth rows of the table. If the target row or column is detected correctly, zero is shown. If the detected row is at the bottom or top of the target row, a positive or negative number is shown respectively. In a similar notation, if the detected column is at the right or left of the target column a positive or negative number is shown respectively. Shaded cells in this table show if there is an error in detection. From total 42 target characters, 27 characters (65%) were detected correctly and 15 characters (35%), shaded in the table, were not detected correctly because their row or column (or both) was identified wrongly. Since no training set was used, low accuracy results were expected. In this study, we paid attention to error cases and tried to find a consistent pattern when they are generated.

TABEL 1

TARGET CHARACTER AND DETECTED CHARACTER IN THE TARGET WORDS PRESENTED TO SUBJECT. THE SHADED CELLS SHOW WHEN AN ERROR OCCURRED. DIFFERENCES BETWEEN COLUMN/ROW OF THE TARGET AND THAT OF THE DETECTED CHARACTERS ARE SHOWN IN THE FOURTH/FIFTH ROW OF THE TABLE, RESPECTIVELY.

WORD	BOWL				CAT			DOG			FISH				FISH				GLOVE				
Target	B	O	W	L	C	A	T	D	O	G	F	I	S	H	F	I	S	H	G	L	O	V	E
Detected	B	O	W	F	C	A	T	D	O	G	F	H	M	H	E	D	S	A	G	H	7	V	E
Column	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	-1	1	0	-1	0	-4	0	0	0
Row	0	0	0	-1	0	0	0	0	0	0	0	0	-1	0	0	-1	0	-1	0	0	3	0	0

Word	HAT			HAT			RAT			SHOES					WATER				
Target	H	A	T	H	A	T	R	A	T	S	H	O	E	S	W	A	T	E	R
Detected	H	B	T	L	A	T	R	A	Z	S	H	A	Q	M	S	A	T	E	R
Column	0	1	0	4	0	0	0	0	0	0	0	-2	0	0	-4	0	0	0	0
Row	0	0	0	0	0	0	0	0	1	0	0	-2	2	-1	0	0	0	0	0

Target Column

Total Error

Target row

												Total Error
												15
	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	0	0	0	0	0	0	0	1
	0	0	0	0	1	3	1	0	0	0	0	5
Target row	0	2	0	0	2	27	1	0	0	1	0	6
	0	0	0	0	0	1	0	0	0	0	0	1
	0	0	0	0	0	1	0	0	0	0	0	1
	0	0	0	0	0	1	0	0	0	0	0	1
	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0
Total Error 15	0	2	0	1	3	6	2	0	0	1	0	

Fig. 3. Error in detecting target row and column as a function of the distance between target and detected row and column. In this figure the target row/column is considered at the center. The shaded cells indicate when an error occurs and the value of the cells show the difference between target and detected row or column.

Fig. 3 is another representation of the row and column differences (the fourth and fifth rows of Table 1). The focus in this representation is how close the detected character is to the target character. It shows that 27 times target character was detected correctly. However, any other nonzero number shows an error in detection (they are shaded cells). In this figure, for example, number “3”, located on the top of 27, indicates that “three times” the detected row was one row above target row while the column was detected correctly or number “2”, located at the far left in the figure and four cells away from 27, indicates that “two times” a column was detected which was on the left side of target column and four columns away from it, while the row was detected correctly. Using the Chebyshev distance, the distance between detected row/column and target row/column was calculated as follows:

$$D = \lim_{k \to \infty} \left(\sum_{i=1}^{2} \left| e_i - c_i \right| \right)^{1/k} = \max_i \left(\left| e_i - c_i \right| \right) \tag{1}$$

where e_i and c_i represent the locations of the target (center) and the wrongly detected characters, respectively. Percentage of the error cases for different values of D is shown in Figure 5.

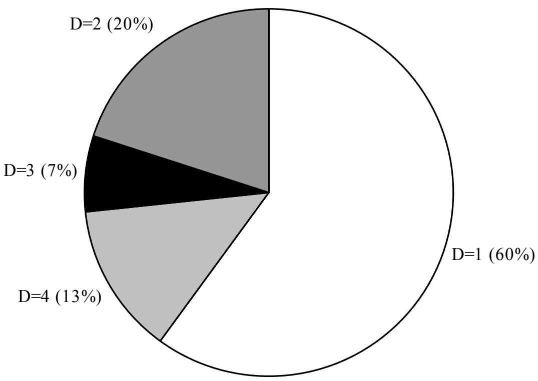

Fig. 4. Percentage of error as a function of distance between target (center cell) and detected row/column. The Chebyshev distance was considered.

From Table 1, Fig. 3 and Fig. 4, it can be concluded that majority of the error cases in detecting a character is when the adjacent row/column to the target row/column was detected (60%). This can be attributed to subject’s attention where the subject is unable to focus precisely only on the target letter and a flashing in adjacent row/column elicits the P300. It should be noted that the number of times that the target row was detected wrongly while the target column was detected correctly is the same as the number of times that the target column was detected wrongly while the target row was detected correctly (6 times). In three cases, the detected column had the distance of four from the target column. In two of these cases, the error in detection was for the character at the beginning of the word (letter “H” in HAT and “W” in WATER). We speculate that in these two cases, the subject was not ready to attend to the target character.

IV. Conclusion

The purpose of this study was to explore perceptual errors in a P300 BCI speller paradigm designed by Farwell and Donchin. In order to reveal the possible sources of error in the paradigm, we did not use any signal processing algorithm that is based on learning such as neural networks. If an algorithm based on learning was used, any systematic error would be learned. Our data analysis showed that majority of error cases happened when an adjacent row or column to the target row/column is detected. Therefore, in the design of any new P300 speller paradigm this perceptual error should be considered. There are two directions for our future work. We will analyze the Farwell and Donchin’s paradigm using more data sets to reveal other possible perceptual sources of errors. In parallel, we will design a new paradigm in which these sources are considered.

References

[1] J. R. Wolpaw and *et al.,* “Brain-computer interfaces for communication and control,” *Clinical Neurophysiology*, vol. 113, pp. 767-791, 2002.

[2] H. H. Jasper, “The ten-twenty electrode system of the international federation,” *Electroencephalogram and Clinical Neurophysiology*, vol. 10, pp. 371-375, 1958

[3] J. R. Wolpaw, D. J. McFarland, G. W. Neat, and C. A. Forneris, “An EEG-based brain-computer interface for cursor control,” *Electroenceph Clin Neurophys*, vol. 78, pp. 252-259, 1991.

[4] G. Pfurtscheller, D. Flotzinger and J. Kalcher, “Brain computer interface-a new communication device for handicapped persons,” *J. of Microcomputer Applications*, Vol. 16, pp. 293-299, 1993.

[5] A. Kostov, and M. Polak, “Parallel man–machine training in development of EEG-based cursor control,” *IEEE Trans. Rehab. Eng.*, vol. 8, pp. 203–205, 2000.

[6] S. A. Hillyard, and M. Kutas, “Electrophysiology of cognitive processing,” *Annual Review of Psychology*, vol. 34, pp. 33–61, 1983.

[7] J. Polich, “P300 in clinical applications,” *Electroencephalography: basic principles, clinical applications and related fields* (Eds. E. Niedermayer and F. Lopes de la Silva). Urban and Schwartzenberger, Baltimore-Munich. pp. 1073-1091, 1999.

[8] L. A. Farwell, and E. Donchin, “Talking off the top of the head: toward a mental prosthesis utilizing event-related brain potentials’, Electroenceph. Clin. Neurophysiol., pp. 510-523, 1988.

[9] J. E. Raymond, K. L. Shapiro, and K. M. Arnell, “Temporary suppression of visual processing in a RVSP task: An attentional blink?” *Journal of Experimental Psychology: Human Perception and Performance*, vol. 18, pp. 849-860, 1992.

[10] N. G. Kanwisher, “Repetition blindness: Type recognition without token individuation,” *Cognition*, vol. 27, pp. 117-143, 1987.

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

[12] D. Ravden, and J. Polich, “On P300 measurement stability: habituation, intra-trial block variation, and ultradian rhythms” *Biological Psychology*, vol. 51, no. 1, pp. 59-76, 1999.

[13] Blankertz, B., and *et al.,* “The BCI competition 2003: progress and perspectives in detection and discrimination of EEG single trials,” *IEEE Trans. Biomed. Eng.*, vol. 51, pp. 1044-52, 2004.

[14] R. Fazel-Rezai, and J. F. Peters, “P300 wave feature Extraction: preliminary results,” *Proceedings of the Canadian Conference of Electrical and Computer Engineering*, Saskatoon, SK, Canada, pp. 376-379, 2005.

[15] R. Fazel-Rezai, and S. Ramanna, “Brain Signals: Feature extraction and classification using rough set methods’, In D. Slezak, J. T. Yao, J. F. Peters, W. Ziarko, X. Hu (Eds.), *Rough Sets, Fuzzy Sets, Data Mining and Granular Computing, Lecture Notes in Artificial Intelligence*, 3642, pp. 709-718, Springer-Verlag, Berlin, 2005.

[16] S. Ramanna, and R. Fazel-Rezai, “P300 Detection based on Rough Sets’, In: J. F. Peters A. Skowron (Eds.), *Transactions on Rough Sets*, 4100, Springer-Verlag, Berlin, pp. 207-227, 2006.