34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

# Reducing Human Error in P300 Speller Paradigm for Brain-Computer Interface

Scott Gavett, *Student Member, IEEE,* Zachary Wygant, *Student Member, IEEE,* Setare Amiri, *Student Member, IEEE,* and Reza Fazel-Rezai, *Senior Member, IEEE*

*Abstract*—**Since the brain-computer interface (BCI) speller paradigm was first introduced by Farwell and Donchin in 1988, there have been many visual modifications to the paradigm. Most of these changes involve the original matrix format such as changes in the number of rows and columns, font size and color, flash time vs. dark time, and flash order. However, recent studies show that there is human error in generating P300 based on this paradigm that none of these changes can help to reduce it. In this study, we analyze this type of error among three paradigms, two based on the matrix structure and one region-based paradigm. It is shown that the human error is reduced significantly in the region-based paradigm.**

## I. Introduction

Brain-computer interface (BCI) spellers are used as a communication tool by people affected with neuromuscular disorders [1]. Diseases such as amyotrophic lateral sclerosis (ALS), brainstem stroke, and brain or spinal cord injuries, to name a few, affect the way a person would normally communicate with the world around them. The virtual keyboard provides people with these communication impairments with a tool to help them communicate with others. BCIs have made significant progress over the last two decades. Substantial research is being conducted to create a consistent and reliable communication channel with the human brain, based on electroencephalography (EEG). One way that the BCI is being implemented uses the P300 potential, a positive peak in the EEG 300 msec after a stimulus. The P300 has been used as a reliable signal for BCI stimulus presentation paradigms [2].

Recently, there has been progress in the improvement of P300 speller accuracy and speed. Various P300 stimuli presentation techniques have been proposed. A typical P300 speller consists of data collection, signal processing, and classification [3]. During the data collection process, a paradigm should be presented to the subject to provoke the P300 potential. Farwell and Donchin proposed the first BCI row-column speller in which a user was presented with a six-by-six matrix of alphanumeric characters [2]. The characters are flashed in random order by rows and columns. When the row or column with the intended character flashes a P300 potential is created. Then the intersection of the row and column with the highest potentials is the intended character. Due to very low amplitude of the P300 in EEG signals, the classification of the P300 requires multiple numbers of flashes to achieve high accuracy.

The Farwell and Donchin paradigm has been quite popular among research groups and has been tested with various configurations and it has been the main paradigm investigated by different research group for the P300 BCI since 1988. Salvaris, et al. investigated modifications in the background color, font size, font style and increasing or decreasing the display area to analyze the classification difference between simple modifications to the visual protocol for the speller [4]. The results showed that although no visual protocol was the best for all subjects, the best performances were obtained with the white background visual protocol and the worst performance was obtained with the small symbol size protocol. Allison further investigated the relationship between the matrix size and EEG measures, detection accuracy and user preferences [5]. Their results indicated that the larger matrices evoked larger P300 amplitudes and the matrix size did not significantly affect the performance or preferences. To further explore that relationship, Guger studied the use of a row-column along with a single character paradigm of the BCI speller over the normal subjects to see the subsequent improvement in the overall accuracy of the system [7]. However, the row-column paradigm provides more accuracy and bit rate as compared to the single character paradigm.

Fazel-Rezai investigated a type of error in the paradigm called “human error” in the P300 BCI [7]. This error is created because of adjacency of different characters in the matrix based P300-speller. It was suggested that research groups should start thinking of new way of presenting characters that is not based on the matrix of characters [7]. To move away from the matrix paradigm, Fazel-Rezai et al. introduced a new paradigm based on grouping characters into regions and detecting the target character in two levels [8]. During his investigation there was an error analysis performed for each region to see how the placement of the

Manuscript received March 28, 2012. This work was supported in part by the NSF funding through ND ESPCoR.

S. Gavett is with the Biomedical Signal Processing Laboratory, Electrical Engineering Department, University of North Dakota, Grand Forks, ND 58202 USA.

Z. Wygant is with the Biomedical Signal Processing Laboratory, Electrical Engineering Department, University of North Dakota, Grand Forks, ND 58202 USA.

Setare Amiri is with the Biomedical Signal Processing Laboratory, Electrical Engineering Department, University of North Dakota, Grand Forks, ND 58202 USA.

R. Fazel-Rezai is with the Biomedical Signal Processing Laboratory, Department of Electrical Engineering, University of North Dakota, Grand Forks, ND 58203 USA (phone: 701-777-3368; fax: 701-777-5253; e-mail: reza.fazel-rezai@engr.und.edu).

regions could affect the results. In this paper, we have shown that the human error existed in the original matrix paradigm is significantly reduced in the region based paradigm.

## II. Methods and Materials

In this paper, we will compare the adjacency problem introduced by human error for the original row-column, single character, and region-based paradigms. The following sections these three paradigms are explained.

### A. *Row/Column paradigm*

Farwell and Donchin introduced the row/column (RC) paradigm [2], most widely discussed and used for P300 BCI. In this paradigm there are a total of 36 characters arranged in a six-by-six matrix as shown in Fig. 1. When the rows and columns flash in a random order, the corresponding row or column with the intended character creates a P300 in EEG signals. The probability of the target being flashed is 0.17 ($\frac{1}{6}$), which is capable of producing a robust P300 [9, 10, 11]. The drawback with such a system is that there is more time required to isolate the targets as more flashes are required. Guger studied the RC for both P300 and motor imagery-based BCI system and discovered that only 72 out of 81 RC subjects spell with an accuracy of 80-100%, while using motor imagery with 99 subjects, only 19% of subjects were able to achieve 80-100% accuracy [6].

The adjacency problem (human error) in this paradigm exists since when the subject focuses on one character but the adjacent characters are flashed and therefore a false P300 is elicited when it should not be generated.

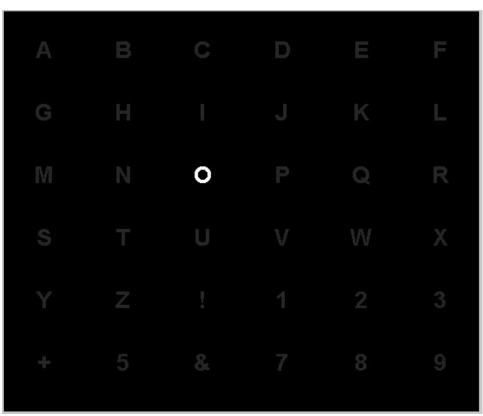


Fig. 1. Row/Column paradigm [2].

### B. *Single Character Paradigm*

For the single character (SC) paradigm, the six-by-six matrix is the same. However, the key difference is that only a single character flashes at a time instead of an entire row or column. Guger compared both the SC and RC paradigms [6] and results suggests that only 55.3% (N=38) were able to spell with 100% accuracy in the SC paradigm as compared to the 72.3% (N=81) of the subjects were able to spell with 100% accuracy in the RC paradigm. This showed that the RC paradigm provided more accuracy than the SC paradigm. Although SC paradigm is totally different than RC paradigm, the same structure (matrix based) exists for this type of paradigm and similarly adjacency problem can lower accuracy due to human error.

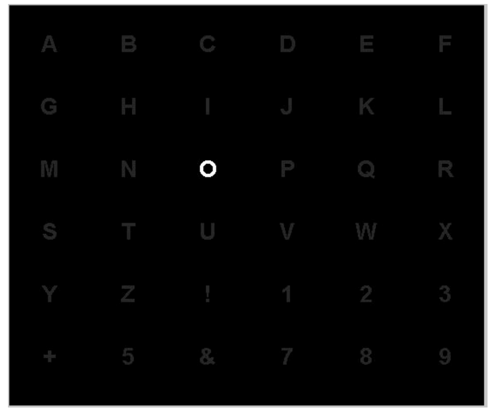


Fig. 2. Single character paradigm [6].

### C. *Region-Based Paradigm*

In the region based (RB) paradigm [8], seven sets of characters are arranged into seven different regions in level 1 as shown in Fig. 3. These regions are then flashed in a random order to generate a P300 potential. After a successful selection of the region, the seven characters will disperse in the same pattern. Also, the individual characters will become larger in size now that there are only seven showing as shown in Fig. 4. The region-based paradigm provides more input character sets, while reducing the crowding effect and adjacency problem.

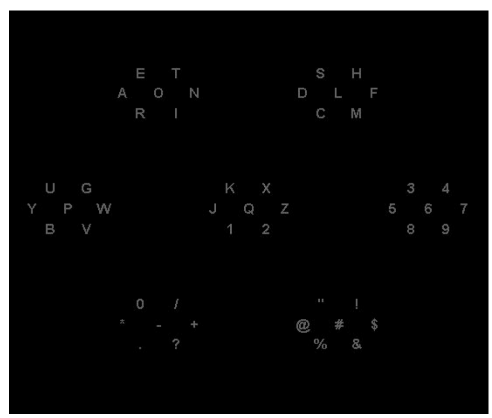


Fig. 3. Region-based paradigm at the first level [8].


Fig. 4. Region-based paradigm at the second level [8].

### *D. Experiments*

The experiments have been approved by the Internal Review Board (IRB-201006-372) at the University of North Dakota. We were able to test all of the paradigms. The tests were run with eight male subjects and 2 female subjects, all between the ages of 20 and 30. For these subjects, we tested the row/column, single character, and the region-based paradigm in a random order.

First, all of the subjects were asked to spell two words, 'WATER' and 'LUCAS' during the calibration of the program. The experiments were performed in random order having three trials for each of the three paradigms. Each subject performed six trials of each paradigm (RC, SC, RB) where they had to spell "PEBBLE!" three times and "MX85+Z&" three times. The reason for choosing these words was to have all regions to be considered as target characters. However, since several characters do not exist in RC and SC paradigms (e.g., ! and &), RC and SC paradigms were modified and unused characters in "PEBBLE!" and "MX85+Z&" such as "Y" and "W" were replaced with "!" and "&" for performing the experiments as shown in Fig. 1 and Fig. 2.

Products of Guger Technologies (g.tec) were used, including g.GAMMAbox and g.USBamp for recording and g.BSanalysis for classification. Six flashes with flash time 100 msec and blank time of 60 msec were considered. EEG signals were recorded from eight channels at FZ, CZ, PZ, OZ, P3, P4, PO7, and PO8 locations as shown in Fig. 5. An electrode at the FPZ location was considered as a ground channel and one electrode on the right mastoid was considered as a reference channel. Data were sampled with a frequency of 256 Hz and filtered by a 0.1 Hz high pass, a 30 Hz low pass. Linear discriminant analysis (LDA) was used for classification.

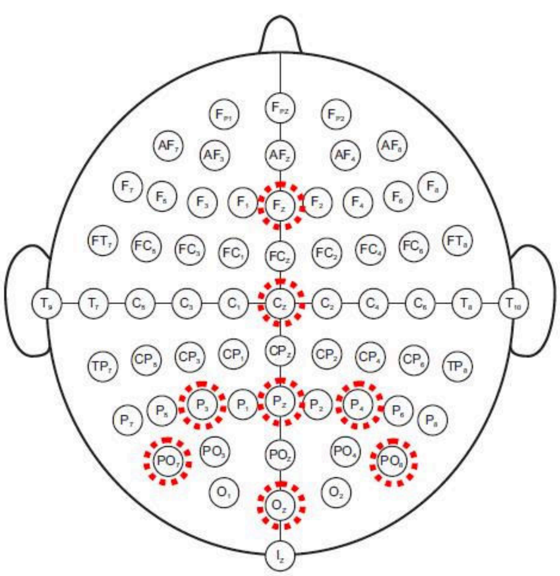

Fig. 5. Placement of electrodes [12]

## III. Results

The human error and adjacency problem was calculated based on the approach that was first introduced in [7] as shown in Fig. 6. In this presentation, the target character is placed in the middle of screen with the number of times that was spelt correctly. The surrounding numbers show the number of errors happened for detecting the target character. Figures 6 and 7 show the errors for paradigms RC and SC, respectively.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | | | | | | 2 | | | | | |
| 4 | | | | | | 2 | | | | | |
| 3 | | | | | | 4 | | 1 | | | |
| 2 | | | | | | 5 | 1 | | | 1 | |
| 1 | 1 | | | | | 8 | | 1 | | | |
| 0 | | 2 | | 6 | 7 | 348 | 5 | 5 | 3 | 1 | 3 |
| -1 | | | | | | 4 | 2 | | 1 | 1 | |
| -2 | | | | | 1 | 1 | | | | | |
| -3 | | | | | | 1 | | | | | |
| -4 | | | | | | 1 | | | | | |
| -5 | | | | | | 2 | | | | | |
| | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 |

Fig. 6. Adjacency errors for the RC paradigm.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | | | | 2 | | 1 | | | | 1 | |
| 4 | | | | | 1 | 4 | 4 | 2 | 1 | | |
| 3 | 1 | | 1 | 2 | 3 | 3 | 3 | 2 | 3 | 4 | |
| 2 | 1 | | 1 | 2 | 2 | 5 | 3 | 4 | 3 | 3 | 1 |
| 1 | 4 | | | 5 | 3 | 11 | 3 | 2 | 2 | | |
| 0 | | 2 | 1 | 5 | 13 | 221 | 3 | 5 | 7 | 1 | 1 |
| -1 | 1 | | 3 | 2 | 3 | 8 | 3 | 1 | 1 | | |
| -2 | 1 | 2 | 2 | 1 | | 3 | 4 | 1 | 1 | | 1 |
| -3 | 1 | | 1 | 1 | 3 | 2 | | | 2 | 1 | |
| -4 | | 1 | 3 | 2 | | | | 1 | | 1 | |
| -5 | | 1 | | | 2 | 3 | 2 | | 3 | | |
| | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 |

Fig. 7. Adjacency errors for the SC paradigm.

Fig. 8 shows the adjacency errors with the RB paradigm. In this presentation, cells with black background show the correct detection of each region while all other numbers indicate the number of errors in detecting the target region.

| Target Region \ Detected Region | Region 1 | Region 2 | Region 3 | Region 4 | Region 5 | Region 6 | Region 7 |
|---|---|---|---|---|---|---|---|
| Region 1 | 96 | 3 | 8 | 5 | 3 | 1 | 4 |
| Region 2 | 6 | 102 | 1 | 5 | 0 | 3 | 3 |
| Region 3 | 7 | 2 | 99 | 5 | 1 | 3 | 3 |
| Region 4 | 6 | 2 | 4 | 94 | 6 | 4 | 4 |
| Region 5 | 4 | 1 | 3 | 1 | 107 | 1 | 3 |
| Region 6 | 3 | 5 | 8 | 4 | 7 | 83 | 10 |
| Region 7 | 3 | 3 | 8 | 8 | 3 | 4 | 91 |
| Total region selected | 125 | 118 | 131 | 122 | 127 | 99 | 118 |
| Wrongfully selected | 29 | 16 | 32 | 28 | 20 | 16 | 27 |
| Region Accuracy | 80.0% | 85.0% | 82.5% | 78.3% | 89.2% | 69.2% | 75.8% |

Fig. 8. Adjacency errors for the RB paradigm. Black cells show the correct detection of each region while all other numbers indicate the number of error in detecting the target region. In each row, the gray cells are adjacent to the black cell in the same row.

In each row, the gray cells are adjacent to the black cells in the same row. Based on the chosen target characters ("PEBBLE!" and "MX85+Z&"), each region was targeted to be selected 120 times. However, they were selected more or less due to errors in selecting target characters as shown in Fig. 8. The number of times that a region was selected incorrectly and the accuracy of each region are shown as well in Fig. 8.

## IV. Discussion

In the RC paradigm there were 72 errors in the detection of the target characters. 33.33% of these errors happened adjacent to the target character. In addition, 86.11% of the error cases happened in the same row or column as the target character. This shows the human error in the RC paradigm that has been investigated and reported previously.

Similar phenomena can be observed in the SC paradigm as shown in Fig. 7. Out of 199 errors, 17.59% of the errors happened in the adjacent characters.

In the RB paradigm, regions are distributed on the computer screen as shown in Fig. 3. Therefore, there is no character adjacency flashing to create a false P300 similar to RC and SC paradigms. Error in seven regions was between 69% for region 6 to 89% for region 5. Although no specific pattern was observed for error in different regions, region error analysis requires data from more subjects that currently is under investigation.

## V. Conclusion

We have shown that there is no adjacency problem in RB paradigm similar to what exists in RC and SC paradigms. Therefore, the adjacency errors that existed in the original matrix-based paradigm are not observed in the RB paradigm. In addition, the accuracy of the RB and RC are statistically better than the SC, with no statistical difference between the RB and RC paradigms. Statistical analysis was performed using a two-way ANOVA test with a post hoc Tukey test.

## References


[1] J. Wolpaw, D. Mcfarland, G. Neat, and C. Forneris, "An EEG-based brain-computer interface for cursor control," *Electroencephalography & Clinical Neurophysiology,* vol. 78, pp. 252-259, 1991.

[2] L. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalography and clinical Neurophysiology,* vol. 70, pp. 510-523, 1988.

[3] 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 Transactions on Rehabilitation Engineering,* vol. 8, pp. 164-173, 2000.

[4] M. Salvaris and F. Sepulveda, "Visual modifications on the P300 speller BCI paradigm," *Journal of Neural Engineering*, vol. 6, doi:10.1088/1741-2560/6/4/046011.

[5] B. Allison, J. Pineda "ERPs evoked by different matrix sizes: Implications for a brain computer interface (BCI) system," *IEEE Transactions on neural systems and rehabilitation engineering*, vol. 11, pp. 110-113, 2003.

[6] C.Guger, S. Daban, E. Sellers, C. Holzner, G. Krausz, R. Carabalona, F. Gramatica, and G. Edlinger "How many people are able to control a P300-based brain-computer interface (BCI)?" *Neuroscience letters*, vol. 462, pp. 94-98, 2009.

[7] R. Fazel-Rezai "Human error in P300 speller paradigm for brain-computer interface," in *29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2007, pp. 2516-2519.

[8] R. Fazel-Rezai, K. Abhari, "A region-based P300 speller for brain-computer interface," *Canadian Journal of Electrical and Computer Engineering*, vol. 34, pp. 81-85, 2009.

[9] J. Polich, "Attention, probability, and task demands as determinants of P300 latency from auditory stimuli," *Electroencephalography and clinical Neurophysiology* vol. 63, pp. 251-259, 1986.

[10] J. Polich, "Task difficulty, probability, and inter-stimulus interval as determinants of P300 from auditory stimuli," *Electroencephalography and Clinical Neurophysiology* vol. 68, pp. 311-320, 1987.

[11] C. Duncan-Johnson, E. Donchin, "The P300 component of the event-related brain potential as an index of information processing," *Biological Psychology*, vol. 14, pp. 1-52, 1982.

[12] F. Sharbrough, C. E. Chatrian, R. P. Lesser, H. Luders, M. Nuwer, T. W. Picton, "American Electroencephalographic Society guidelines for standard electrode position nomenclature," *Journal for Clinical Neurophysiology*, vol. 8, pp. 200-202, 1991.