Neuroscience Letters 531 (2012) 63–68

Contents lists available at SciVerse ScienceDirect

Neuroscience Letters

journal homepage: www.elsevier.com/locate/neulet

Plenary article

A general P300 brain–computer interface presentation paradigm based on performance guided constraints

George Townsend[a], Jessica Shanahan[a], David B. Ryan[b], Eric W. Sellers[b,*]

[a] Algoma University, Sault Ste. Marie, Ontario P6A 2G4, Canada
[b] East Tennessee State University, 807 University Parkway, Johnson, TN 37614, USA

ARTICLE INFO

Article history:
Received 14 May 2012
Received in revised form 6 July 2012
Accepted 22 August 2012

Keywords:
Brain–computer interface
EEG
P300
Event-related potential
Rehabilitation
Amyotrophic lateral sclerosis

ABSTRACT

An electroencephalographic-based brain–computer interface (BCI) can provide a non-muscular method of communication. A general model for P300-based BCI stimulus presentations is introduced – the "m choose n" or $C(m$ (number of flashes per sequence), n (number of flashes per item)) paradigm, which is a universal extension of the previously reported checkerboard paradigm (CBP). $C(m,n)$ captures all possible (unconstrained) ways to flash target items, and then applies constraints to enhance ERP's produced by attended matrix items. We explore a $C(36,5)$ instance of $C(m,n)$ called the "five flash paradigm" (FFP) and compare its performance to the CBP. Eight subjects were tested in each paradigm, counter-balanced. Twelve minutes of calibration data were used as input to a stepwise linear discriminant analysis to derive classification coefficients used for online classification. Accuracy was consistently high for FFP (88%) and CBP (90%); information transfer rate was significantly higher for the FFP (63 bpm) than the CBP (48 bpm). The $C(m,n)$ is a novel and effective general strategy for organizing stimulus groups. Appropriate choices for "m," "n," and specific constraints can improve presentation paradigms by adjusting the parameters in a subject specific manner. This may be especially important for people with neuromuscular disabilities.

1. Introduction

Brain–computer interfaces (BCIs) can reestablish communication for people whose motor and communicative abilities have been compromised by neuromuscular disease [24]. In amyotrophic lateral sclerosis (ALS) motor neuron death eventually paralyzes the patient. BCI technology does not require neuromuscular activity and can help restore rudimentary communication. Introduced by Farwell and Donchin [6], the P300 based BCI uses electroencephalography (EEG) to recognize the P300 response to an attended stimulus. The P300 is an event-related potential (ERP) consisting of a positive deflection in the EEG over parietal cortex occurring approximately 300 ms after the presentation of a rare ('oddball') meaningful stimulus [5].

In the row/column paradigm (RCP) [6], a 6×6 matrix of characters appear on-screen and subjects attend to the item they wish to select. The matrix rows and columns flash randomly and (in theory) flashes containing the attended item will elicit a P300 [5]. Townsend et al. [21] introduced the checkerboard paradigm (CBP) to overcome specific issues that lead to errors in the RCP. First, incorrect selections are typically adjacent to the intended item [3,7]. To overcome this problem the CBP presents flash groups in pseudo-random order instead of rows and columns. Second, the order in which rows and columns flash is unconstrained allowing the same item to flash multiple times in succession. Multiple flashes of an attended item can change the ERP in a deleterious manner. A second flash of the same item may go unnoticed, which will not result in a P300, and the target epochs overlap temporally, reducing P300 amplitude or changing ERP morphology [15,23]. To overcome this problem, the CBP places a minimum of six intervening flashes between subsequent flashes of any given matrix item.

Many different stimulus organization and presentation methods have been introduced and shown to mitigate errors associated with the RCP [10,15,17,19,20]. The CBP has shown that changing the organization of the stimulus presentation increases the probability of an attended item being classified correctly. Fig. 1 shows discriminant values assigned to each cell of the matrix. The attended item was "W." The CBP is shown in the left column of the figure and the RCP is shown in the right column of the figure. In both cases, a SWLDA classifier was trained and then tested online (scores normalized). The particular character and the subject were both chosen at random from the Townsend et al. [21] dataset. Comparing the discriminant values for the rows and columns including the target item (W) showed that the mean discriminant value of 14.35 in the CBP condition was significantly lower than the mean discriminant value of 28.78 in the RCP condition ($t_{16} = 3.81$, $p = .002$). These data

* Corresponding author. Tel.: +1 423 439 4476; fax: +1 423 439 5695.
E-mail address: sellers@etsu.edu (E.W. Sellers).

0304-3940/$ – see front matter © 2012 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.neulet.2012.08.041

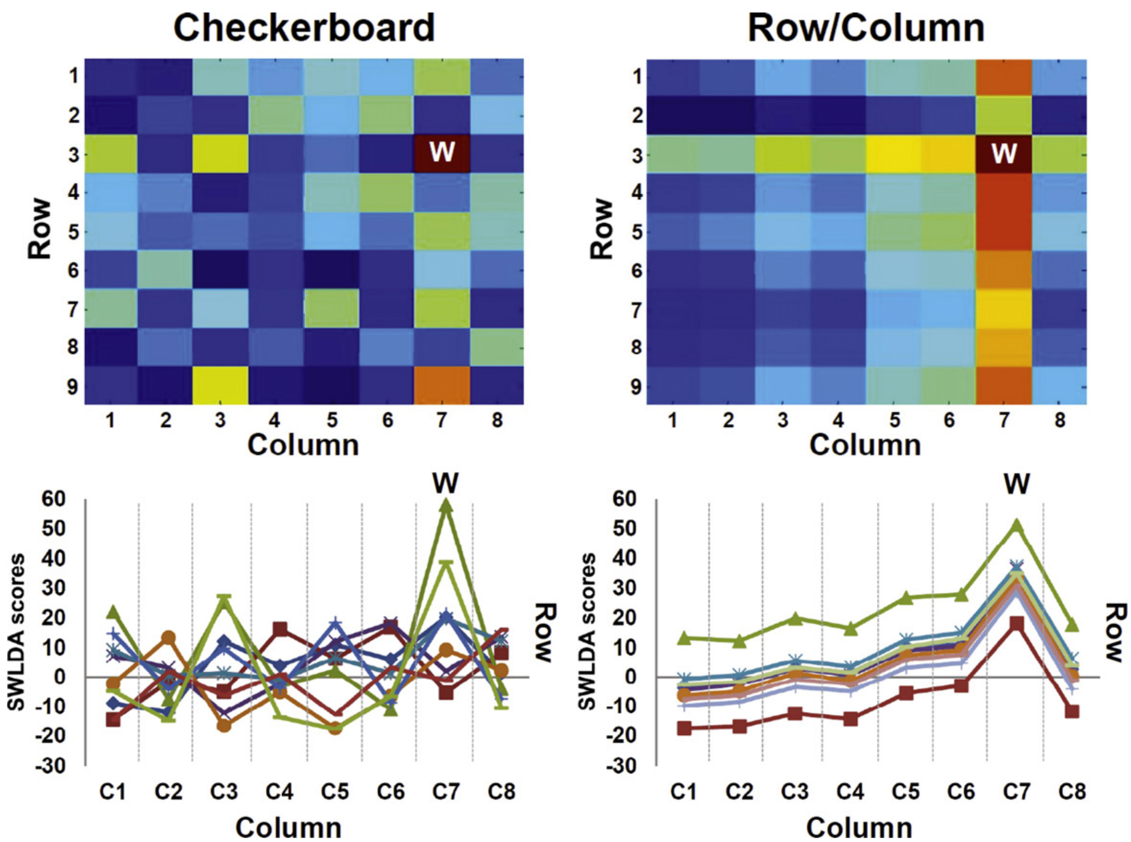

Fig. 1. Top row: eight column × nine row plot of SWLDA scores for each cell. Blue = low scores; red = high scores. Intersection of column seven and row three indicate the desired item (i.e. "*W*"). Checkerboard shown on left; row/column shown on right. Bottom row: SWLDA scores for plotted by each row and column for the Checkerboard paradigm (left) and SWLDA scores for the row/column paradigm (right). Although the attended item was correctly chosen by the classifier in both paradigms, the discriminant values are clearly higher in the RCP for the entire row (horizontal dark green lines) and column (column 7) of the attended character. In contrast, the CBP discriminant values are much lower overall, and cell locations having higher values are not obviously related to the attended item.

show that the CBP is less likely to produce an incorrect response, as compared to the RCP.

Other paradigms placing constraints on when and how matrix items flash have also produced increases in speed and accuracy. Jin et al. [12] varied the number of items that flash simultaneously. The goal was to use the minimum number of flashes to uniquely identify all of the stimuli in the matrix. Results showed that optimizing stimulus group size for each subject improved performance. Allison and Pineda [1] examined the effects of flashing multiple rows or columns simultaneously. Martens and Leiva [16] described a generative model-based method of stimulus presentation and classification.

1.1. *C(m,n): a universal extension of the checkerboard paradigm*

The "*m* choose *n*" paradigm is a universal extension of the CBP that captures all possible ways to present stimuli [22]. It is a general approach and it is important to note that the paradigm can generate the RCP, the CBP, or any other implementation. Here we define a *stimulus group* as the number of items that flash simultaneously. Stimulus groups are defined in mathematical sets independent of their location in the matrix. *M* is the number of stimulus groups and a *sequence* is completed once all m have flashed once. *N* is the number of times individual matrix items flash. An *item pattern* is the temporal presentation order of each matrix item. For example, the Farwell and Donchin 6 × 6 paradigm conforms to a $C(12,2)$ implementation. There are a total of 12 flashes per sequence and each matrix item is presented two times per sequence (i.e., each of the six rows, and each of the six columns flash one time per sequence). The Farwell and Donchin paradigm employed the following constraints: (1) each of the twelve stimulus groups are arranged in the rows and columns that appear in the matrix; (2) stimuli are presented sequentially at random without replacement; and (3) once the twelve stimulus groups have been presented, presentation order is re-randomized and the process is repeated until a criterion is met (e.g., the desired number of sequences has been presented).

Jin et al. [13] proposed a paradigm similar to $C(m,n)$, based on combinations computed with binomial coefficients. However, the method reported in [13] was unconstrained. In a follow-up paper, Jin et al. [11] extended their approach to constrain double item flashes and adjacent item flashes; the same constraints we reported in [21]. The present work goes further to introduce a formal and general approach that can derive any stimulus presentation arrangement. After all possible unconstrained ways to flash stimuli have been identified, constraints designed to enhance ERP discrimination are applied because of the exceedingly large number of possible stimulus groups. The specific constraints employed here and the choices for "*m*" and "*n*" have been selected because some have previously shown to improve BCI performance, and all of them logically conform to the basic principles of the Oddball paradigm [5].

1.2. *Present study: the five flash paradigm (FFP)*

The present study compares the CBP to a particular implementation of $C(m,n)$ called the five flash paradigm (FFP). The CBP conforms to $C(24,2)$ where in a sequence of 24 stimulus groups each matrix item flashes two times. $C(36,5)$ was chosen to exploit

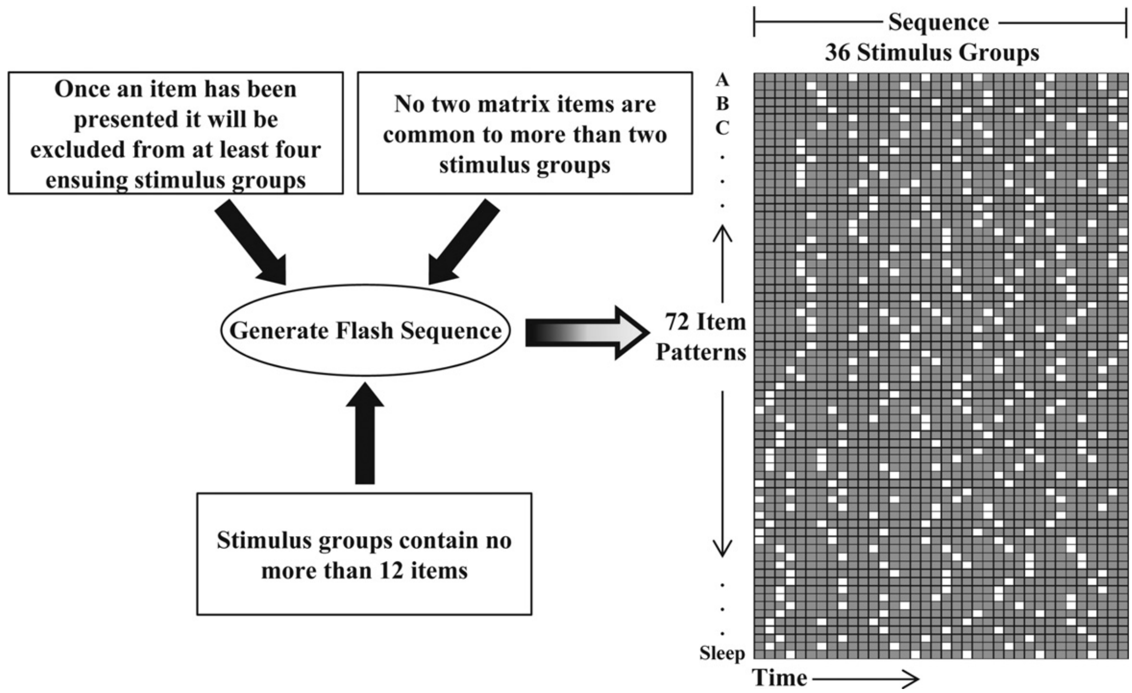

Fig. 2. The constraints employed in the *C*(36,5) method and an example stimulus sequence. Left: the three performance-based constraints imposed to generate a sequence of flashes. Right: an example sequence based on the constraints. White cells represent flashing matrix items. For each of the 36 stimulus groups (columns) no more than twelve items are included in any stimulus group. For each item pattern (rows), item-specific flashes are separated by a minimum of four non-flash events, and no target shares more than two flashes in common with any other. Therefore any three of the five flashes are sufficient to uniquely identify the target. The minimum spacing between flashes as well as the wrap-around from the end of a sequence to the start of another sequence can also be confirmed by examining the figure.

improvements realized by the CBP (as compared to RCP) and further improve speed and accuracy. The FFP adds three flashes of each matrix item with 12 additional stimulus groups (60% more matrix items with 33% additional sequences). Thus, we hypothesize that item selections in the FFP will be made more rapidly because fewer sequences should be required.

Unconstrained, there are *C*(36,5) or 372,992 possible candidates from which a set of 72 must be chosen:

$$\binom{372992}{72} = \frac{372992 \times 372991 \times 372990 \times \ldots \times 372921}{72 \times 71 \times 70 \times \ldots \times 1} \qquad (1)$$

This leads to a large number of possible candidates, on the order of 2×10^{297}.

The FFP enforces three constraints to mitigate problems associated with other stimulus presentation paradigms (e.g., RCP and CBP). Different constraints may be necessary to accommodate different matrix sizes and SOAs. To construct each sequence, the FFP simultaneously applies the following three constraints:

(1) *Ensure that once an item has been presented it will be excluded from at least four subsequent stimulus groups*. A similar constraint is used in the CBP. Subjects may not notice sequential flashes of attended items. In this case, a P300 will not result. In addition, excluding a given item for a minimum of four subsequent stimulus groups ensures non-overlapping target epochs. Overlapping epochs can reduce P300 amplitude or change ERP morphology [15,23].
(2) *Ensure that no two matrix items are common to more than two stimulus groups*. This allows any matrix item to be uniquely identified with three flashes, allowing for unique identification of items even if two target items are not perceived. In the CBP, only two flashes of each item are contained in one sequence. Thus, if either target item flash is not perceived, the target item cannot be uniquely identify. The problem is exacerbated in the RCP (see Fig. 1) because each matrix item is shared with an entire row and an entire column. Thus, in the RCP it is not possible to uniquely identify target items unless all flashes of the attended item in a sequence generate appropriate ERP responses.
(3) *Require that stimulus groups contain no more than 12 items*. This constraint provides large enough stimulus groups for faster presentation than the CBP, which has a fixed stimulus group size of six items. Adding up to twice the number of stimuli to each stimulus group results in shorter sequence duration.

At random an initial stimulus group is selected thereby reducing the possible number of viable candidates. The second stimulus group is randomly selected and evaluated; if the group conforms to the constraints it will be included. Otherwise, it is rejected and another stimulus group is selected. This process continues until a sequence of 72 viable candidates have been selected or all possible stimulus groups have been sampled. If all possible candidates have been evaluated before a complete sequence has been identified, the algorithm will restart and continue to iterate until a completed sequence (of 72 viable candidates) is found. Fig. 2 depicts an example set of stimulus groups for the FFP. We hypothesize that the FFP will produce superior performance than the CBP because it presents many more target stimuli per unit of time at the expense of a small increase in sequence length (or time taken per sequence).

2. Materials and methods

2.1. Subjects

Eight subjects participated in the study (five men, three women). Subjects were recruited from the Algoma University community. None had uncorrected visual impairments or any known cognitive deficit. The study was approved by the Algoma University board of ethics and subjects gave informed consent.

Table 1
The CBP and FFP are compared. The effective bitrates exclude the 3.5 s pause between selections; the practical bitrates includes the 3.5 s pause between selections. The difference in accuracy was not statistically significant whereas the bitrate increase was statistically significant.

Subject	Checkerboard paradigm						Five flash paradigm					
	Flashes	Acc	Practical		Effective		Flashes	Acc	Practical		Effective	
			Sel (min)	Bitrate	Sel (min)	Bitrate			Sel (min)	Bitrate	Sel (min)	Bitrate
1	48	100	5.45	33.65	10	61.70	36	100	6.32	38.97	13.33	82.27
2	48	92	5.02	30.96	9.20	56.76	36	100	6.32	38.97	13.33	82.27
3	120	72	2.16	13.33	2.88	17.77	72	64	2.74	16.92	4.27	26.33
4	48	92	5.02	30.96	9.20	56.76	72	80	3.43	21.15	5.33	32.91
5	48	96	5.24	32.31	9.60	59.23	36	88	5.56	34.29	11.73	72.39
6	48	88	4.80	29.62	8.80	54.30	36	96	6.06	37.41	12.80	78.98
7	72	96	4.11	25.38	6.40	39.49	36	76	4.80	29.62	10.13	62.52
8	72	100	4.29	26.44	6.67	41.13	36	84	5.31	32.73	11.20	69.10
Mean	63	92	4.51	27.83	7.84	48.39	45	88	5.07	31.26	10.27	63.34

2.2. Data acquisition, processing

Subjects sat in a chair approximately 1 m from a computer monitor that displayed the 8 × 9 matrix. EEG was recorded with an 8-channel electrode cap with tin electrodes (Electro-Cap International). All channels were referenced to the right mastoid and grounded to the left mastoid. An 8-channel monopolar g.tec (Guger Technologies) amplifier was used to acquire the EEG (amplification to ±2 V before ADC; high-pass 0.1 Hz; digitization rate 256 Hz; notch filter at 60 Hz). The eight electrode sites (based on the 10/20 international electrode placement system) were Fz, Cz, P3, Pz, P4, P07, P08 and Oz [14]. BCI2000 software [18] controlled stimulus presentation, data collection and online processing.

2.3. Experimental paradigm

Subjects completed two sessions on different days within a one-week period. Sessions were counterbalanced; half of the subjects began with CBP and half with the FFP. Each session consisted of a calibration phase and an online test phase. In the calibration phase, subjects were shown a word to spell. Each letter of the word was the target or attended item in serial order. The word was displayed at the top left corner of the screen with the current target in parenthesis (“PHONE” (P)). Subjects were instructed to attend to each flash of the target item and to repeat this procedure for each letter of the word, as prompted. Subjects were given six words, totaling 30 letters, during the calibration phase. The online phase consisted of five different words, also containing five characters.

2.4. Classification

As described in [14], independent SWLDA classifiers were derived for the CBP and FFP [4]. In the CBP and FFP calibration phases, each item selection included the same number of item flashes, consistent with the approach used in [21]. That is, each matrix item was presented 10 times before the classifier chose an item. In the FFP this was achieved by presenting two sets of 36 stimulus groups (each item flashed five times in each of two sequences). In the CBP this was achieved by presenting five sets of 24 stimulus groups (each item flashed two times in each of the five sequences). Thus, both paradigms used 300 target flashes to acquire training data. We used the SWLDA algorithm to determine the signal features that best discriminated between target and non-target flashes (MATLAB version 7.6 R2007a, stepwise fit function).

During online classification, epochs from each stimulus item were averaged before applying the SWLDA classification coefficients. In both paradigms, the coefficients were applied to the specific spatiotemporal features of each of the 72 items and summed. The item with the highest score was selected and presented to the subject as feedback.

2.5. Determining the optimal number of sequences

Due to the relatively low signal-to-noise ratio of the P300, it is necessary to flash each item multiple times and average the resulting epochs [2]. During calibration, the number of target flashes was constant across subjects and presentation methods. During the online testing phase, the number of flashes used online was optimized according to each subject’s maximum written symbol rate (WSR, or symbols/min; [8]) based on the training data.

3. Results

3.1. Online accuracy and bit rate

Table 1 contains the number of sequences, accuracy, selections/min and bitrate for each subject and paradigm. Ultimately, the performance of any BCI must be measured by its performance in an online implementation. It is important to note that only online results are reported here. In this study, online bit rate in the FFP paradigm was 63 bits per minute (bpm), 31% higher than the 48 bpm observed in the CBP paradigm ($\sigma = 0.04$). The difference in online accuracy (CBP 92%; FFP 88%) was not statistically significant ($\sigma = 0.10$).

3.2. Waveform morphologies

Fig. 3 shows target responses for each of the eight subjects averaged across two groups of electrodes. The right panel shows Fz, Cz, Pz, P3, and P4. The left panel shows Oz, PO7, and PO8. The electrodes were grouped in this way because within each group morphology is very similar and across each electrode group morphology is notably different. As can be seen in the figure, the waveforms for each subject are generally representative of the grand means shown in the bottom row of Fig. 3. It is also notable that the FFP and CBP produce very similar morphology for all subjects except for subject 2 and 8, and this is only true in the averaged electrodes shown in the left column.

4. Discussion

The first goal of this study was to formalize the concept of flashing targets organized in a matrix to capture all possible (unconstrained) organizations and to provide a description of how constraints designed to enhance the performance of the interface are implemented. The second goal of the study was to explore a particular implementation of this concept by introducing the five

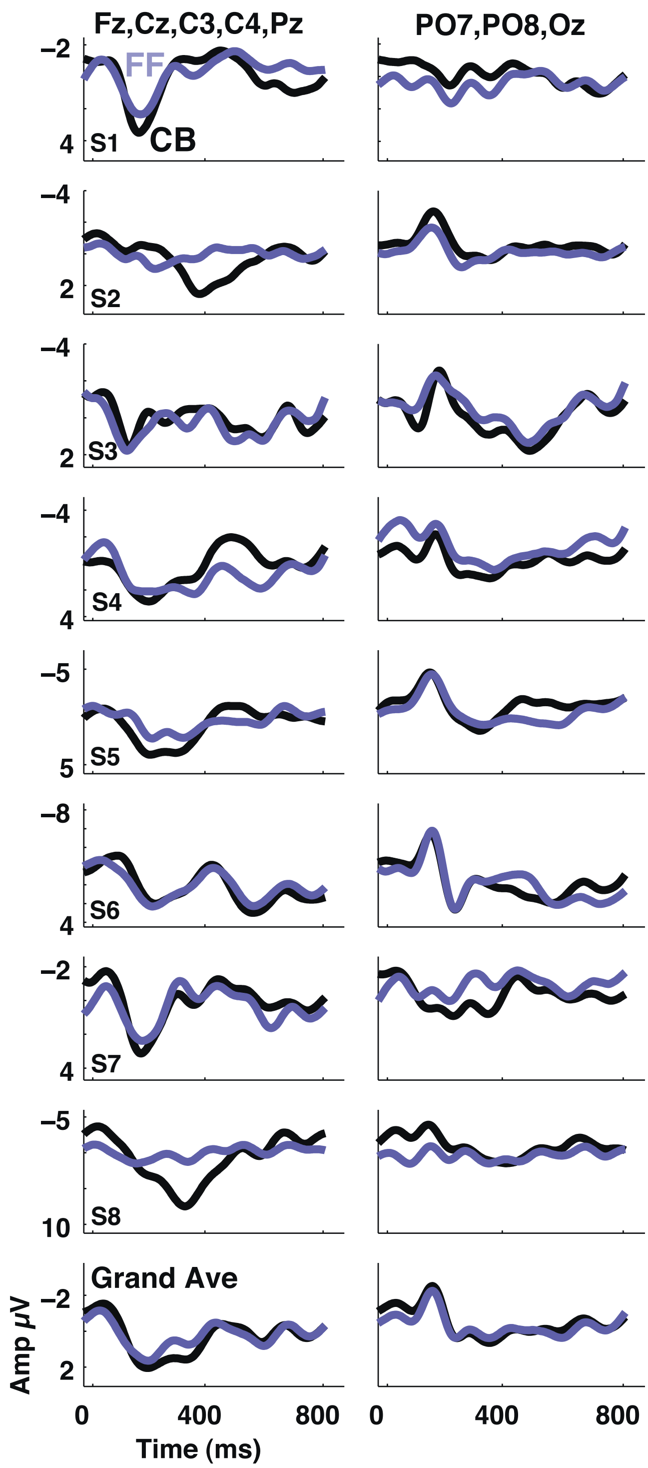

Fig. 3. Waveform morphologies of each subjects target responses for the FFP (blue) and CBP (black) are shown by electrode set. The bottom row shows grand averages.

flash paradigm (FFP) and compare it to the checkerboard paradigm (CBP). Many researchers have developed different ways to flash targets in a matrix; however, each approach is uniquely described. In this study we have generalized these approaches with the aim of generating a better understanding of what organizations are possible and what is required to make them viable. The $C(m,n)$ method allows rows and columns of the matrix to be completely disassociated, unless the constraints are chosen to enforce a row/column stimulus presentation. Therefore, it provides a formal model of previous studies involving stimulus group organization as well as yet to be described paradigms.

Both paradigms achieved high accuracy and bit rates. With either paradigm, the P300 BCI was calibrated in approximately 12 min, similar to the results reported in [9]. As in the earlier CBP study [21], the 9×8 matrix was used. The matrix emulates most functions provided by a standard keyboard, thereby producing an ecologically valid BCI that can be used for a variety of tasks. The FFP increased bitrate by using fewer flashes to produce an optimal level of performance. Moreover, the improvement in bitrate was realized despite the fact significant differences in accuracy between the two paradigms were not detected.

5. Conclusions

Brain–computer interfaces provide severely disabled people a communication option that does not depend on neuromuscular control. Improved BCI performance should subsequently improve quality of life for people with severe motor disabilities. Slow, inaccurate systems can be frustrating to use and cause deleterious effect. This study shows the importance of carefully designing presentation paradigms for P300 BCIs. The development of the CBP was a significant step in improving the P300 interface and the $C(m,n)$ model has expanded upon these improvements. Additional work (to be reported elsewhere) has begun to focus on variations of the model. Preliminary data using a $C(20,3)$ variant has shown online performance resulting in a mean bitrate of 133 bpm ($n = 3$) excluding time between selections. The best subject was 97% accurate with a bitrate of 143 bpm excluding time between selections (which is 48 bpm including time between selections). Thus, additional generalizations of the model should continue to improve the speed and accuracy of BCI systems.

Acknowledgements

We thank our financial supporters NIH/NIBIB & NINDS (EB00856); NIH/NIDCD (R21 DC010470-01); NOHFC (990015); FEDNOR (842-504726); Algoma University. Special thanks to the New York State Department of Health's Wadsworth research centre for their generous support to the Algoma University BCI lab.

References

[1] B.Z. Allison, J.A. Pineda, Effects of SOA and flash pattern manipulations on ERPs, performance, and preference: implications for a BCI system, International Journal of Psychophysiology 59 (2006) 127–140.
[2] J. Cohen, J. Polich, On the number of trials needed for P300, International Journal of Psychophysiology 25 (1997) 249–255.
[3] E. Donchin, K.M. Spencer, R. Wijesinghe, The mental prosthesis: assessing the speed of a P300-based brain–computer interface, IEEE Transactions on Rehabilitation Engineering 8 (2000) 174–179.
[4] N.R. Draper, H. Smith, Applied Regression Analysis, Wiley, New York, 1981, pp. xiv 709.
[5] M. Fabiani, G. Gratton, D. Karis, E. Donchin, Definition, identification, and reliability of measurement of the P300 component of the event related potential, Advances in Psychophysiology 2 (1987) 1–78.
[6] L.A. Farwell, E. Donchin, Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials, Electroencephalography and Clinical Neurophysiology 70 (1988) 510–523.
[7] R. Fazel-Rezai, Human error in P300 speller paradigm for brain–computer interface, Conference Proceedings: IEEE Engineering in Medicine and Biology Society (2007) 2516–2519, 2007.
[8] A. Furdea, S. Halder, D.J. Krusienski, D. Bross, F. Nijboer, N. Birbaumer, A. Kubler, An auditory oddball (P300) spelling system for brain–computer interfaces, Psychophysiology 46 (2009) 617–625.
[9] C. Guger, S. Daban, E. Sellers, C. Holzner, G. Krausz, R. Carabalona, F. Gramatica, G. Edlinger, How many people are able to control a P300-based brain–computer interface (BCI)? Neuroscience Letters 462 (2009) 94–98.
[10] B. Hong, F. Guo, T. Liu, X. Gao, S. Gao, N200-speller using motion-onset visual response, Clinical Neurophysiology 120 (2009) 1658–1666.
[11] J. Jin, B.Z. Allison, E.W. Sellers, C. Brunner, P. Horki, X. Wang, C. Neuper, An adaptive P300-based control system, Journal of Neural Engineering 8 (2011) 036006.

[12] J. Jin, B.Z. Allison, E.W. Sellers, C. Brunner, P. Horki, X. Wang, C. Neuper, Optimized stimulus presentation patterns for an event-related potential EEG-based brain–computer interface, Medical & Biological Engineering & Computing 49 (2011) 181–191.

[13] J. Jin, P. Horki, C. Brunner, X. Wang, C. Neuper, G. Pfurtscheller, A new P300 stimulus presentation pattern for EEG-based spelling systems, Biomedizinische Technik Biomedical Engineering 55 (2010) 203–210.

[14] D.J. Krusienski, E.W. Sellers, D.J. McFarland, T.M. Vaughan, J.R. Wolpaw, Toward enhanced P300 speller performance, Journal of Neuroscience Methods 167 (2008) 15–21.

[15] S.M. Martens, N.J. Hill, J. Farquhar, B. Scholkopf, Overlap and refractory effects in a brain–computer interface speller based on the visual P300 event-related potential, Journal of Neural Engineering 6 (2009) 026003.

[16] S.M. Martens, J.M. Leiva, A generative model approach for decoding in the visual event-related potential-based brain–computer interface speller, Journal of Neural Engineering 7 (2010) 26003.

[17] M. Salvaris, F. Sepulveda, Visual modifications on the P300 speller BCI paradigm, Journal of Neural Engineering 6 (2009) 046011.

[18] G. Schalk, D.J. McFarland, T. Hinterberger, N. Birbaumer, J.R. Wolpaw, BCI2000: a general-purpose brain–computer interface (BCI) system, IEEE Transactions on Biomedical Engineering 51 (2004) 1034–1043.

[19] E.W. Sellers, D.J. Krusienski, D.J. McFarland, T.M. Vaughan, J.R. Wolpaw, A P300 event-related potential brain–computer interface (BCI): the effects of matrix size and inter stimulus interval on performance, Biological Psychology 73 (2006) 242–252.

[20] K. Takano, T. Komatsu, N. Hata, Y. Nakajima, K. Kansaku, Visual stimuli for the P300 brain–computer interface: a comparison of white/gray and green/blue flicker matrices, Clinical Neurophysiology (2009).

[21] G. Townsend, B.K. LaPallo, C.B. Boulay, D.J. Krusienski, G.E. Frye, C.K. Hauser, N.E. Schwartz, T.M. Vaughan, J.R. Wolpaw, E.W. Sellers, A novel P300-based brain–computer interface stimulus presentation paradigm: moving beyond rows and columns, Clinical Neurophysiology 121 (2010) 1109–1120.

[22] G. Townsend, J. Shanahan, G.E. Frye, E.W. Sellers, P300-BCI: Disassociating Flash Groups from Physical Organizations Provides Improved Performance. Presented at the Fourth International Brain–Computer Interface Meeting, Asilomar, CA, June, 2010.

[23] M.G. Woldorff, Distortion of ERP averages due to overlap from temporally adjacent ERPs: analysis and correction, Psychophysiology 30 (1993) 98–119.

[24] J.R. Wolpaw, N. Birbaumer, D.J. McFarland, G. Pfurtscheller, T.M. Vaughan, Brain–computer interfaces for communication and control, Clinical Neurophysiology 113 (2002) 767–791.