An Assistive Communication Brain–Computer Interface for Chinese Text Input

James W. Minett, Gang Peng, Lin Zhou, Hong-Ying Zheng, William S-Y. Wang

Language Engineering Laboratory
Department of Electronic Engineering
The Chinese University of Hong Kong
Hong Kong SAR, China

***Abstract*—The performance of assistive communication brain–computer interfaces has been studied mostly for languages with alphabetic script. The viability of using such systems for languages with other types of script, such as Chinese, which has a logographic script, is currently poorly understood. Here, a performance analysis of the P300 Speller is presented for Chinese text input. The performance of six distinct paradigms, based on the established Row/Column (RC) and Single Character (SC) spellers, are tested and compared for 30 able-bodied, native Chinese readers. In terms of accuracy per trial, the optimal paradigm is based on the SC speller: 63.3% of participants were able to achieve 80% or better classification accuracy for 15 trials. However, because the RC speller has shorter trial duration than the SC speller, the optimal paradigm in terms of communication rate is a variant of the RC speller in which stimuli are intensified by changing background color. A communication rate of 14.5 bits per minute was attained using this paradigm. For a lexicon of ~11,000 Chinese characters, this corresponds to a projected mean input rate of ~1.1 characters per minute.**

Keywords-Assistive communication; electroencephalography (EEG); brain–computer interface (BCI); P300 Speller; Chinese

I. INTRODUCTION

Most people have the ability to communicate with others, whether by speaking, by writing, by signing, or otherwise. However, a substantial minority of people experience nervous system damage that limits their ability to control voluntary muscle movement, in some cases preventing them from communicating by any conventional means. Various types of brain–computer interface (BCI) have been developed that harness scalp-recorded electroencephalograph (EEG) signals to assist communication. EEG signals that have been used for this purpose include the P300 event-related potential [1–3], slow cortical potentials [4], and sensorimotor rhythms [3, 5].

The most commonly used method is the P300 Speller, developed by Farwell and Donchin [1]. In its original formulation, a matrix of 6 × 6 alphanumeric characters and control sequences is presented to the user, the rows and columns of the matrix being intensified in random order. The user inputs a character (the *target*) by attending to it, keeping a running count of the number of times it is intensified. Because intensification of the target is a rare event (probability 1/6) in an oddball sequence [6], it elicits a P300 response with greater magnitude than non-targets, allowing the user's target character to be determined. The P300 Speller has been shown to be easier for users to control accurately than sensorimotor-based systems [3].

Various studies have examined the performance of variants of the P300 Speller. For example, in [7], reducing matrix size from 6 × 6 to 3 × 3 tended to increase online classification accuracy, but at the expense of reduced communication rate. A trend was also observed that increasing the inter stimulus interval from 175 ms to 350 ms decreased performance in terms of both accuracy and communication rate [7]. In addition, stimulus intensity has been shown to influence the amplitude of the P300 elicited, with visually more salient intensification giving rise to greater classification accuracy [8].

Much of the research on assistive communication BCIs has focused on the input of English and other languages that are written using alphabetic script. Few studies have examined the efficacy of BCIs for inputting languages that use other types of script, such as Chinese, which has a logographic script comprising more than 11,000 characters [9]. Chinese characters

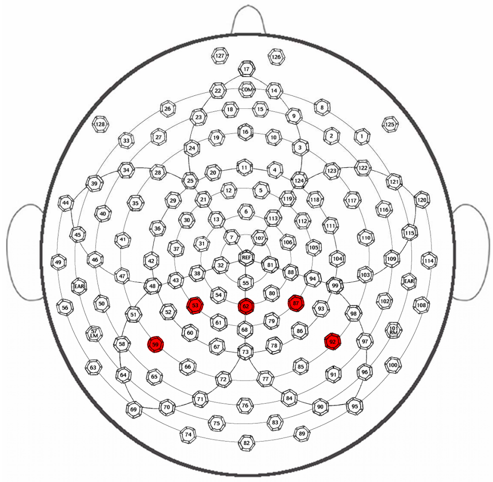

Figure 1. Electrode locations (Pz, P3, P4, P7, and P8, highlighted in red; viewed from above the head) used for EEG measurements.

The work described in this paper was supported by research grants awarded to W. S-Y. W. by the Innovation and Technology Commission, The Government of the Hong Kong SAR (Grant No.: ITS/068/09) and the Shun Hing Institute of Advanced Engineering at The Chinese University of Hong Kong (Grant No.: BME 8115020).

一	丨	丿	乀	乙	紅
孖	彳	儿	才	八	力
打	口	厂	白	中	王
火	言	日	鉄	寺	草
們	江	木	天	大	女
二	十	石	怵	山	←

Figure 2. Matrix of components used in the experiment.

are made up of hierarchical arrangements of strokes. A P300-based BCI has been developed that allows users to input Chinese characters stroke-by-stroke [10]. However, Chinese characters may consist of 20 or more strokes, limiting the efficiency of such a system when applied to the entire lexicon of ~11,000 characters. A more efficient method is sought by which Chinese characters can be input by making a small number of selections of subcharacter components, comprising individual strokes, stroke patterns, and radicals [11, 12].

II. THE EXPERIMENT

The purpose of the present experiment is to test the extent to which native Chinese readers are able to control a P300-based BCI to select Chinese characters and subcharacter components—comprising individual strokes, stroke patterns, and radicals—and to determine which of six distinct stimulus presentation paradigms provides the greatest accuracy and communication rate in so doing.

A. Participants

Thirty native Chinese readers (14 female, 16 male, age: 21.7 ± 2.8), all students at The Chinese University of Hong Kong, were paid to participate in this study. No participant had any prior experience with BCIs. All were able-bodied, had normal or corrected-to-normal vision, and reported no history of neurological illness. Each participant gave informed consent. Approval to conduct the study was obtained from the Survey and Behavioural Research Ethics Committee of The Chinese University of Hong Kong.

B. Equipment and Data Acquisition

Participants were seated in a quiet room in front of a LCD screen. EEG data were acquired from five electrode locations (Pz, P3, P4, P7, P8; see Fig. 1) using a Net Amps 200 amplifier (Electrical Geodesics, Inc.) and 128-channel Ag/AgCl electrode arrays. EEG data were recorded continuously at a rate of 1,000 Hz, referenced to the vertex, filtered with an analogue band-pass filter (0.1 Hz to 400 Hz), and digitized using a 16-bit A/D converter. The data were recomputed offline against average-mastoid reference, and low-pass filtered at 40 Hz. EEG segments for critical stimuli were extracted from 100 ms before stimulus onset to 400 ms after stimulus onset, the mean

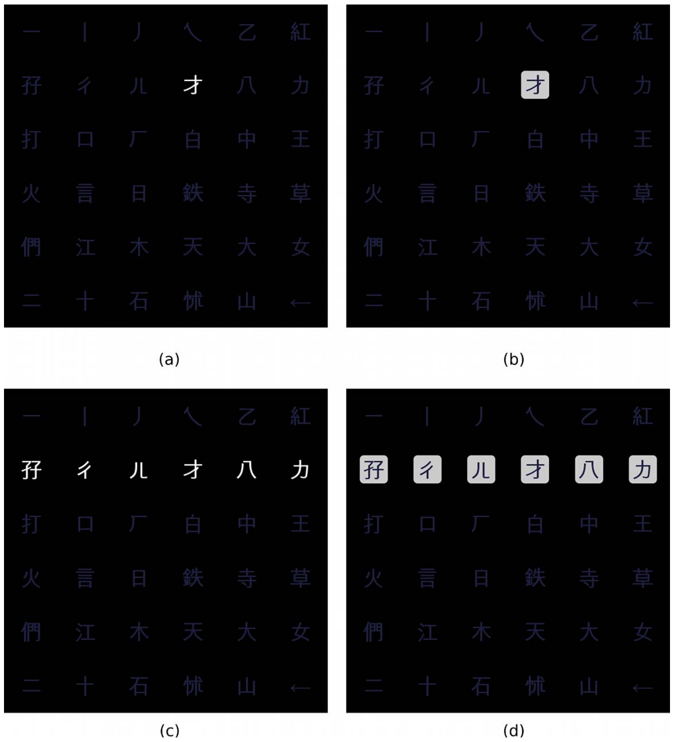

Figure 3. (a) Paradigm I intensifies a single component by changing foreground color. (b) Paradigm II intensifies a single component by changing background color. (c) Paradigms III & IV intensify a whole row or column of components by changing foreground color. (d) Paradigms V & VI intensify a whole row or column of components by changing background color.

voltage in the 100 ms interval prior to stimulus onset being used for baseline correction. No artifact correction was applied to deal with eye blinks or eye movements.

C. Procedure

The experiment consisted of six blocks, each using a distinct paradigm for stimulus presentation. The same set of stimuli were used for each paradigm: a 6×6 matrix of thirty-six components (see Fig. 2) consisting of thirty-five Chinese characters, radicals, stroke patterns, or strokes, all extracted from the Traditional character set used in Hong Kong, plus one control sequence (‘←’). The components were rendered on-screen using the Songti font, and subtended a visual angle of 0.70^{O} H × 0.70^{O} W. The participant’s task was to attend to a specific component of the matrix, shown in advance on-screen, keeping a running count of the number of times it was intensified.

Two paradigms were based on the Single Character (SC) speller [3], in which a single component was intensified at a time, either by changing the foreground color, i.e., text color (Paradigm I, illustrated in Fig. 3(a)), or by changing the background color (Paradigm II, illustrated in Fig. 3(b)). Four paradigms were based on Farwell and Donchin’s Row/Column (RC) speller [1], in which a whole row or column of components was intensified at a time, either by changing the foreground color (Paradigms III & IV, illustrated in Fig. 3(c)), or by changing the background color (Paradigms V & VI, illustrated in Fig. 3(d)). The stimulus onset asynchrony was set to 117 ms for Paradigms I, II, III and V,

TABLE I. SUMMARY OF PRESENTATION PARADIGMS

Paradigm	Intensification Selection	Intensification Method	Stimulus Onset Asynchrony	Trial Duration
I	single component	foreground	117 ms	4.2 s
II	single component	background	117 ms	4.2 s
III	row / column	foreground	117 ms	1.4 s
IV	row / column	foreground	167 ms	2.0 s
V	row / column	background	117 ms	1.4 s
VI	row / column	background	167 ms	2.0 s

and to 167 ms for Paradigms IV and VI — the stimulus itself was displayed for 100 ms, followed by a blank mask of either 17 ms or 67 ms. During each trial, every character (Paradigms I & II) or row/column (Paradigms III–VI) was intensified once. Table I summarizes the parameter settings of each paradigm.

Participants undertook a practice run of 8–12 trials for each paradigm. After the practice runs were completed, the six blocks were conducted in random order. Each block consisted of 5 runs, each comprising 18 trials with the same target. The EEG signals obtained from the first two runs of each block were used as training data for classification. The EEG signals obtained from the final three runs of each block were used as testing data for offline analysis of classification accuracy and communication rate. Participants were provided no online feedback regarding classification accuracy.

D. Classification

Fisher's linear discriminant analysis (FDA) was used for offline classification of testing data. FDA operates by seeking a discriminant vector that separates two or more classes optimally; see [2] for an overview of the FDA procedure. The optimal discriminant vector was calculated for each participant and paradigm using feature vectors obtained from the first two runs of the paradigm, concatenating the EEG data collected at the five electrodes, Pz, P3, P4, P7 and P8, from 50 ms to 400 ms after onset of intensification of each component. Classification accuracies for 1–20 trials were estimated by Monte Carlo simulation, drawing 1,000 pseudo-random samples (with replacement) from the testing data for each participant–paradigm combination.

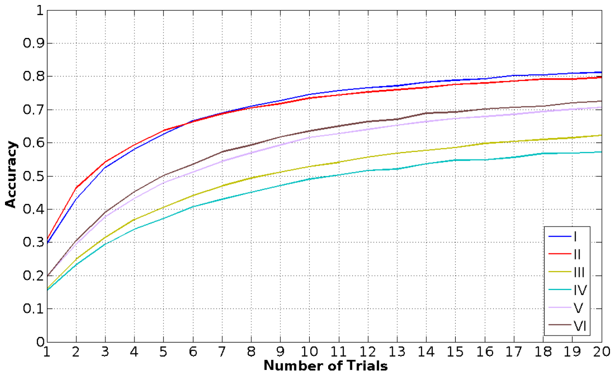

Figure 4. Mean classification accuracy of each paradigm as a function of number of trials.

III. RESULTS

The thirty participants completed a total of 175 blocks. The mean classification accuracy for each paradigm is summarized in Fig. 4. The two SC-based paradigms (I & II) attained greater accuracy than the four RC-based paradigms (III–VI) for all numbers of trials. For the RC-based paradigms, background intensification (Paradigms V & VI) attained greater accuracy than foreground intensification (Paradigms III & IV) for all numbers of trials. The extent of individual difference in classification accuracy across participants is exemplified for Paradigm V in Fig. 5, which plots the accuracy attained by each participant who completed that block.

For classification based on 5 trials (Table II), 43.3% of participants achieved *high* (i.e., 80% or better) accuracy using Paradigm II (with background intensification), 10.0% more than achieved the same accuracy using Paradigm I (with foreground intensification). No more than 13.8% of participants were able to achieve high accuracy using any of the four RC-based paradigms. Focusing on the paradigm with which each participant achieved the greatest accuracy (refer to the column labeled "Best" in Table II), 50% of participants were able to achieve high accuracy; 76.7% were able to achieve 60% or better accuracy. Mean accuracy using each participant's best paradigm was 75.7%.

For classification based on 15 trials (Table III), 63.3% of participants achieved high accuracy with both Paradigm I and II. For the RC-based paradigms, best accuracy was achieved using background intensification (Paradigms V & VI), with more than 40% of participants attaining high accuracy. 80% of participants were able to achieve high accuracy with one of the six paradigms; 93.3% were able to achieve 60% or better accuracy. Mean accuracy using each participant's best paradigm was 90.1%. In terms of classification accuracy, therefore, Paradigms I and II, based on the SC speller, provide the best performance.

For the six paradigms studied, the durations of trials differ across paradigms. This interacts with classification accuracy to influence communication rate. Fig. 6 shows that Paradigm V provides the greatest communication rate, with a peak rate of 14.5 bits per minute, calculated using the formula given in [13]. This paradigm is based on the RC speller, using background intensification and short stimulus onset asynchrony (117 ms).

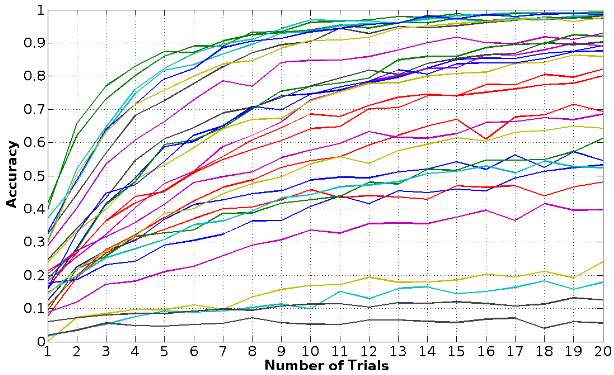

Figure 5. Variation in classification accuracy attained by participants for Paradigm V (one curve per participant).

TABLE II. ESTIMATED PERCENTAGE OF PARTICIPANTS THAT ACHIEVED SPECIFIED CLASSIFICATION ACCURACY FOR 5 TRIALS USING EACH PARADIGM.

Classification accuracy %	SC Speller		RC Speller				Best
	I	*II*	*III*	*IV*	*V*	*VI*	
80 – 100	33.3	43.3	3.3	6.9	13.8	11.1	*50.0*
60 – 79	30.0	13.3	16.7	17.2	17.2	29.6	*26.7*
40 – 59	16.7	16.7	40.0	13.8	27.6	29.6	*10.0*
20 – 39	6.7	13.3	13.3	41.4	27.6	14.8	*13.3*
0 – 19	13.3	13.3	26.7	20.7	13.8	14.8	*0.0*
Mean Accuracy	62.7	63.6	40.6	37.3	47.9	50.1	*75.7*

TABLE III. ESTIMATED PERCENTAGE OF PARTICIPANTS THAT ACHIEVED SPECIFIED CLASSIFICATION ACCURACY FOR 15 TRIALS USING EACH PARADIGM.

Classification accuracy %	SC Speller		RC Speller				Best
	I	*II*	*III*	*IV*	*V*	*VI*	
80 – 100	63.3	63.3	26.7	24.1	48.3	40.7	*80.0*
60 – 79	13.3	16.7	33.3	20.7	17.2	29.6	*13.3*
40 – 59	6.7	3.3	10.0	17.2	17.2	14.8	*6.7*
20 – 39	13.3	13.3	16.7	24.1	3.4	3.7	*0.0*
0 – 19	3.3	3.3	13.3	13.8	13.8	11.1	*0.0*
Mean Accuracy	78.9	77.5	58.5	54.8	67.3	69.2	*90.1*

When the stimulus onset asynchrony is increased to 167 ms, the peak communication rate drops to 10.8 bits per minute. A similar drop in communication rate is also observed for the RC speller with foreground intensification (Paradigms III & IV). Paradigms I and II achieve a peak communication rate of 10.0 bits per minute. In terms of communication rate, therefore, Paradigm V, a variant of Farwell and Donchin's original Row/Column speller in which rows and columns are intensified by changing background color, provides the best performance.

IV. CONCLUSION

This study indicates that the P300 Speller is a viable assistive communication BCI to input Chinese characters by selecting subcharacter components. Using a variant of the Row/Column speller in which stimuli are intensified by changing background color, a peak communication rate of 14.5 bits per minute was achieved. A practical BCI for Chinese character input should make it possible to input ~11,000 extant Chinese characters, i.e., about 13½ bits. The results presented here demonstrate that a projected mean communication rate of about 1.1 characters per minute for the full lexicon is feasible. As Sellers and colleagues remark regarding the target users of such assistive communication BCI systems, "when one has virtually no way to communicate, even one selection a minute can make a significant difference in the quality of life" [13].

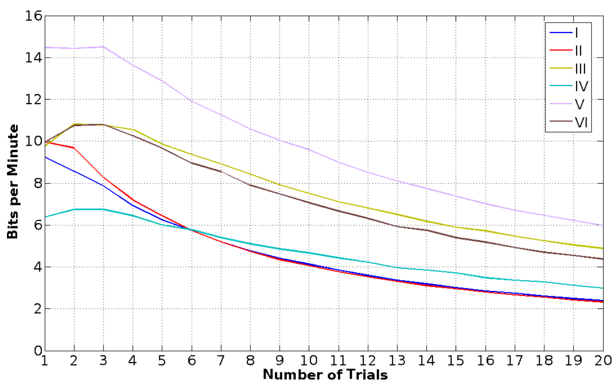

Figure 6. Mean communication rate in bits per minute of each paradigm as a function of number of trials.

REFERENCES

[1] L. A. Farwell, and E. Donchin, "Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials," Electroencephalogr. Clin. Neurophysiol., vol. 70, pp. 510–23, 1988.

[2] U. Hoffmann, J-M. Vesin, T. Ebrahimi, and K. Diserens, "An efficient P300-based brain–computer interface for disabled subjects," J. Neurosci. Methods, vol. 167, pp. 115–125, 2008.

[3] 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)?," Neurosci. Lett., vol. 462, pp. 94–98, 2009.

[4] N. Birbaumer, N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey, A. Kübler, J. Perelmouter, E. Taub, and H. Flor, "A spelling device for the paralysed," Nature, vol. 398, pp. 297–298, 1999.

[5] J. R. Wolpaw, and D. J. McFarland, "Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans," Proc. Natl. Acad. Sci. U. S. A., vol. 101, pp. 17849–17854, 2004.

[6] M. Fabiani, G. Gratton, D. Karis, and E. Donchin, "Definition, identification, and reliability of measurement of the P300 component of the event-related brain potential," Adv. Psychophysiol., vol. 2, pp. 1–78, 1987.

[7] E. W. Sellers, D. J. Krusienski, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "A P300 event-related potential brain-computer interface (BCI): The effects of matrix size and inter stimulus interval on performance," Biol. Psychol., vol. 73, pp. 242–52, 2006.

[8] Z-W. Ma and S-K. Gao, "P300-based brain–computer interface: Effect of stimulus intensity on performance," J. Tsinghua Univ. (Sci. & Tech.), vol. 48, no. 3, pp. 415–418, 2008 [in Chinese].

[9] Xinhua Zidian, Xianggang: Shang wu yin shu guan Xianggang, 1989. [新華字典. 香港商務印書館].

[10] B. Wu, Y. Su, J-H. Zhang, X. Li, J-C. Zhang, W-D. Cheng, and X-X. Zheng, "A virtual Chinese keyboard BCI system based on P300 potentials," Acta Electronica Sinica, vol. 37, no. 8, pp. 1733–1745, 2009 [in Chinese].

[11] Y. P. Chen, D. A. Allport, and J. C. Marshall, "What are the functional orthographic units in Chinese word recognition: The stroke or the stroke pattern?" Q. J. Exp. Psychol. A, vol. 49, no. 4, pp. 1024–1043, 1996.

[12] S-L. Yeh and J-L. Li, "Sublexical processing in visual recognition of Chinese characters: Evidence from repetition blindness for subcharacter components," Brain Lang., vol. 88, pp. 47–53, 2004.

[13] E. W. Sellers, A. Kübler, and E. Donchin, "Brain–computer interface research at the University of South Florida Cognitive Psychology Laboratory: the P300 Speller," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 14, no. 2, pp. 221–224, 2006.