'Write' but not 'spell' Chinese characters with a BCI-controlled robot*

Jin Han[1], *Student Member, IEEE*, Minpeng Xu[1,2], *Member, IEEE*, Yijun Wang[3], *Member, IEEE*, Jiabei Tang[1], Miao Liu[1,2], *Member, IEEE,* Xingwei An[1,2], *Member, IEEE,* Tzyy-Ping Jung[1,4], *Fellow, IEEE,* and Dong Ming[1,2], *Senior Member, IEEE*

***Abstract*— Visual brain-computer interface (BCI) systems have made tremendous process in recent years. It has been demonstrated to perform well in spelling words. However, different from spelling English words in one-dimension sequences, Chinese characters are often written in a two-dimensional structure. Previous studies had never investigated how to use BCI to 'write' but not 'spell' Chinese characters. This study developed an innovative BCI-controlled robot for writing Chinese characters. The BCI system contained 108 commands displayed in a 9*12 array. A pixel-based writing method was proposed to map the starting point and ending point of each stroke of Chinese characters to the array. Connecting the starting and ending points for each stroke can make up any Chinese character. The large command set was encoded by the hybrid P300 and SSVEP features efficiently, in which each output needed only 1s of EEG data. The task-related component analysis was used to decode the combined features. Five subjects participated in this study and achieved an average accuracy of 87.23% and a maximal accuracy of 100%. The corresponding information transfer rate was 56.85 bits/min and 71.10 bits/min, respectively. The BCI-controlled robotic arm could write a Chinese character '福' with 16 strokes within 5.7 seconds for the best subject. The demo video can be found at https://www.youtube.com/watch?v=A1w-e2dBGl0. The study results demonstrated that the proposed BCI-controlled robot is efficient for writing ideogram (e.g. Chinese characters) and phonogram (e.g. English letter), leading to broad prospects for real-world applications of BCIs.**

I. Introduction

Brain-computer interfaces (BCIs) provide an alternative approach for people to directly communicate with the outward environment, which is independent or less dependent of the muscles and peripheral nerves [1-5]. Among all BCI paradigms, the P300 speller [6] and the steady-state visual evoked potential (SSVEP)-based BCI speller [7] are the two most popular vision-based BCI systems.

Recently, spelling with vision-based BCIs has made tremendous progress and gained increasing attention from researchers. Townsend and Platsko developed an asynchronous paradigm for a P300 speller, which boosted the information transfer rate (ITR) beyond 100 bits/min [8]. Nakanishi et al. first applied a task-related component analysis (TRCA) method to detect targets in a 40-target SSVEP-based BCI, which achieved an unprecedented ITR up to 376.58 bits/min [9]. However, most of these studies were focused on spelling English alphanumeric characters and no study investigated how to use a BCI to 'write' as opposed to 'spell' Chinese characters. In 2010, Minett et al. applied the sub-character component method to a row/column (RC) P300 paradigm to spell Chinese character, which reached a peak ITR of 14.5 bits/min [10]. Later, Minett et al. use the shape-based method to encode more than 7000 Chinese characters and achieved an input speed of one character per 107s [11]. In 2016， Yu et al. developed the Hanyu Pinyin-based method and achieved an averaged offline accuracy of 92.6% with a mean ITR of 39.2 bits/min [12]. Although there are many methods to encode Chinese characters, all of these studies are "spell" rather than really "write" Chinese characters.

One reason of this problem is that Chinese characters have a large number and complex structure with two-dimensional planar topology. The other reason is the limitation of BCI technologies (e.g. instruction set, the consuming time for outputting a command). To overcome the internal obstacles of Chinese characters, we first proposed a pixel-based writing method that mapped the starting point and ending point of each stroke of Chinese characters to BCI commands (also known as pixels). And we further proposed a hybrid P300-SSVEP BCI-controlled robot system with an ever-largest instruction set of 108 commands, which could write arbitrary Chinese characters.

II. Materials and Methods

A. Subjects

Five healthy volunteers (two females and three males, 22-25 years of age; all right-handed) with normal or corrected to normal vision participated in both offline and online experiments. The experimental procedure was approved by the Institutional Review Board at Tianjin University. All of the subjects were fully informed of the experiment procedures and signed an informed consent agreement, in accordance with the Declaration of Helsinki, and including a

*Research supported by National Key Research and Development Program of China (No. 2017YFB1300302), National Natural Science Foundation of China (No. 81630051, 61976152, 81601565), Young Elite Scientist Sponsorship Program by CAST (2018QNRC001) and Tianjin Key Technology R&D Program (No. 17ZXRGGX00020). (Jin Han and Minpeng Xu contributed equally to this work.) (Corresponding authors: Yijun Wang; Dong Ming. e-mail: wangyj@semi.ac.cn; richardming@tju.edu.cn).

[1]College of Precision Instruments and Optoelectronics Engineering, Tianjin University, Tianjin 300072, China.

[2]Academy of Medical Engineering and Translational Medicine, Tianjin University, Tianjin 300072, China.

[3]The State Key Laboratory on Integrated Optoelectronics, Institute of Semiconductors, Chinese Academy of Sciences, Beijing 100083, China.

[4]Swartz Center for Computational Neuroscience, University of California, San Diego, CA 92093, USA

statement that they have known all possible consequences of the study.

B. System Description

The proposed BCI-controlled robot system comprises two main subsystems: the hybrid P300-SSVEP BCI system and the robot control system. The robot arm was used to realize the control application of the hybrid BCI system in real environment. Fig. 1 shows the overall control schematic of the BCI-controlled robot system. The hybrid P300-SSVEP BCI system consisted of raw signal recording, pre-processing and classification method. The hybrid P300-SSVEP BCI communicated with the robot control system via TCP/IP communication protocol. The output results of the hybrid BCI system will first be presented on the screen in the form of visual feedback. Only when the results are correct can the robot be moved within 4 seconds, otherwise the participants can cancel the command by gritting their teeth within 4 seconds. In order to widen the applicability of the system, for those who cannot implement gritting their teeth, the cancel command can be replaced by other behaviors, such as continuous blinking.

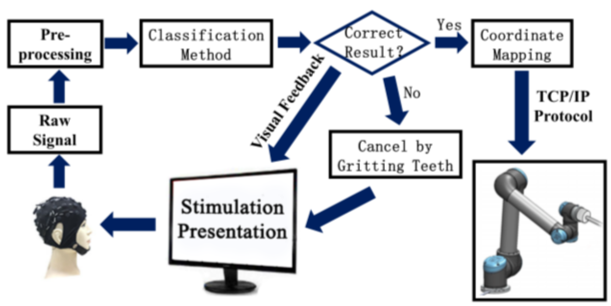

Figure 1. Block diagram of the proposed BCI-controlled robot system.

C. Pixel-based writing method

To overcome the challenges of writing Chinese characters, we proposed a pixel-based writing method. As we all know, the basic element of Chinese characters is strokes. Strokes with different distributions make up different Chinese characters. Based on this, the stimulation interface with 108 commands of the hybrid P300-SSVEP BCI system could be seen as a writing panel. So the main problem was converted into how to write strokes on the stimulation interface. The core of the method was that the starting point and ending point of each stroke of Chinese characters were mapped to a stimulation interface with 108 targets (also called pixels). Connecting the starting point and ending point makes up each stroke of a Chinese character.

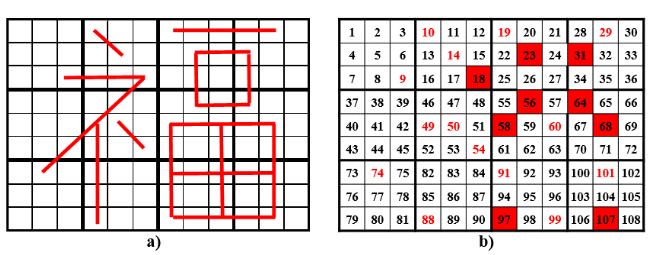

Figure 2. Illustration of Chinese character “福”. a) The two-dimensional pixels of “福” on the stimulation interface; b) The **target** pixels of “福”. The red pixels indicated that the participant needed to select the pixel once as the target, and the pixel with a red filled background indicated that the participant needed to select it twice.

To further illustrate this method, the Chinese character “福” was explained as an example. The starting and ending points of each stroke were regarded as the target pixels. The BCI system wrote “福” in the following sequence: 10-14-9-18-18-74-49-88-50-54-19-29-23-56-23-31-31-64-56-64-58-97-58-68-68-107-91-101-60-99-91-107 (see Fig. 2). It can be seen that the writing method used the odd-numbered endpoints of the sequence as the starting point of the stroke, and the even-numbered endpoints as the ending point of the stroke. Connecting the starting point and the ending point made up all the strokes of the Chinese character “福”.

D. A Hybrid P300-SSVEP BCI Paradigm

This study proposed a hybrid BCI paradigm that embeds the steady-state visual stimuli into the oddball paradigm. A 9×12 matrix showed 108 numbers on a white background (see Figure 3a). They were further divided into 12 small 3×3 matrices. Each small matrix was an independent P300 sub-speller whose characters were individually highlighted by a gray square in a random and ergodic sequence. Each stimulation pixel subtended 1.49 degrees of visual angle in the vertical direction and 1.78 degrees in the horizontal direction. The stimulus duration for each pixel was 200ms and the inter-stimulus interval (ISI) was -100ms. All sub-spellers were triggered at the same time. Therefore, it needed only 1 second to run a complete cycle for all the pixels, which was defined as a ‘round’ in this study. Different from the traditional P300 paradigm, the stimulation pixel changed its grayscale in a sinusoidal mode whose frequency and initial phase were different for each sub-speller, as shown in Figures 3b, 3c, and 3d.

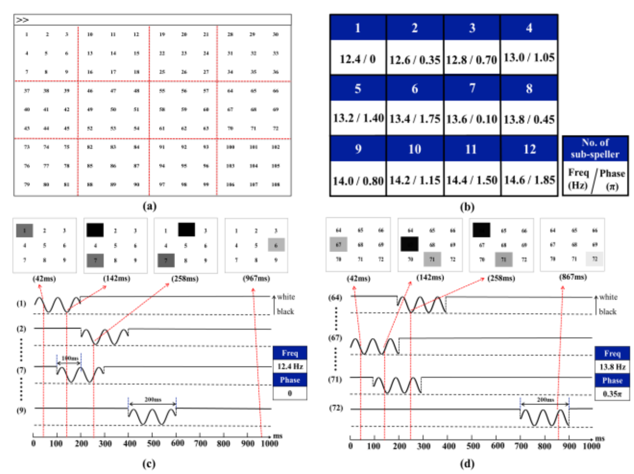

Figure 3. The hybrid P300-SSVEP paradigm. (a) The layout of the 108 pixels on the screen was divided into 12 sub-spellers by the red dash lines. (b) The selected frequency and initial phase of stimulation squares were displayed for each sub-speller. (c) Stimulation process for sub-speller 1. The red dotted lines with arrows indicate specific time points. (d) Stimulation process for sub-speller 8.

To separate the SSVEP frequency band from the P300 frequency band, the 12 flickering frequencies were selected above 12 Hz, from 12.4 to 14.6 Hz with a step of 0.2 Hz. The initial phases were optimized by a search of phase interval (from 0 to 2π with a step of 0.05π) on a public SSVEP dataset using a simulation method [13]. The visual stimuli were presented on a 27-inch liquid-crystal display (LCD) monitor whose resolution was 1,920×1,080 pixels and the refresh rate was 120 Hz. The stimulation program was developed under MATLAB (MathWorks, Inc.) using the Psychophysics Toolbox Version 3.

E. Robot Control System

The UR10 (Universal Robots, Inc.), a 6-axis collaborative robot, was adopted in this study. The absolute maximum distance that the UR10 could reach was 1300mm from the center of the base. It's worth noting that the UR10 can be directly controlled by the hybrid BCI system in real time.

The position of the robot is specified by the 6D coordinates (X, Y, Z, RX, RY, RZ). For the calibrated operating system, X represents the height from the console. Y and Z represent the vertical and horizontal axes of the console, respectively. RX, RY, RZ represent the rotation angles of X, Y, Z axes, respectively.

F. Experimental Procedure

During the experiment, the participants were seated in front of a monitor screen at a distance of 60 cm. They were asked to pay overt attention to a specified pixel beforehand and silently count the number of times the target pixel was intensified.

In the offline experiment, the pixel specified for selection would be indicated by a red triangle underneath with 0.79 degrees of visual angle for 0.7 seconds. Then the visual stimulus ran for five successive rounds for all the pixels, which last 5 seconds, i.e. the subject chose the same target five times. Each round contained one target stimulus and 8 non-target stimuli. All subjects were required to choose all 108 pixels as targets on the screen sequentially, which were divided into three blocks (36 pixels in each block). They would have a break of several minutes between two successive blocks. The offline experiment lasted about 13 minutes for each subject. The offline data were used to analyze the system and determine the parameters of the online experiment, so the robot control is not included.

In the online experiments, all participants were asked to write a Chinese character "福" with 16 strokes (i.e. 32 starting points and ending points) using the proposed BCI-controlled robot system at least 32 times. It is worth noting that the moving time of robot control system was 4s after receiving a command from the hybrid BCI system. Only one round was used for the online spelling test.

G. EEG Recording and Pre-processing

EEG signals were recorded by a Neuroscan Synamps2 system with eight electrodes placed at Fz, Cz, Pz, PO7, PO8, O1, O2, and Oz according to the International 10/20 system. The reference electrode was put on the left mastoid and the ground electrode was placed on the prefrontal lobe. The recorded signals were bandpass filtered at 0.1–200 Hz and the notch filter was set to 50 Hz, digitized at 1000 Hz and then stored on a computer.

In the recognition process, there were two sequential steps: (1) recognizing the sub-speller containing the target pixel and then (2) recognizing the target pixel within the identified sub-speller, as shown in Fig. 4.

Based on our work [14], we used both the P300 and SSVEP features for target recognition. The 8-channel EEG signals (Fz, Cz, Pz, PO7, PO8, O1, O2, and Oz) were filtered by a filter bank (including eight Chebyshev Type I filters) into [X Hz, 92 Hz] (X=1, 11, 22, 34, 46, 58, 70 and 82), and then down-sampled to 250 Hz. The hybrid features were extracted from 50ms to 450ms. The outputs indicated the predicted sub-speller or target pixel.

Figure 4. The total flow diagram of signal processing.

H. Ensemble Task-Related Component Analysis (TRCA)

The Ensemble TRCA has been proved the most powerful recognition algorithm for SSVEP classification [9]. TRCA is an algorithm that maximizes the covariance of task-related components between trials. In the process of calculation, the best projection direction was obtained by defining restriction and using Lagrange multiplier method. The filter bank method could be combined to extract more features and improve the classification performance. Finally, the correlation coefficient between the projection of the test data and averaged individual template was calculated to predict the target pixel. For more details, please refer to [9].

I. Coordinate Mapping

The stimulation interface is a 2-D coordinate system (length* width: m*n), while the robot control system is a 6-D coordinate system. Therefore, after the output results of the hybrid BCI system, the following steps are required to convert the coordinate systems. After the operating platform of the robot control system (length*width: M*N) was calibrated, the position coordinates (x, y) on the stimulation interface correspond to the 6D coordinates of the robot system as $(X, Y-\frac{y}{n-1}*N, Z+\frac{x}{m-1}*M, RX, RY, RZ)$, where X, RX, RY, RZ are fixed constants after the robot system was calibrated.

J. Performance Evaluation

This study uses classification accuracy and ITR as evaluation indicators, which have been widely adopted in BCI research. The ITR can be calculated as [15]:

$$\mathrm{ITR} = \{log_2\mathrm{N} + \mathrm{P}log_2\mathrm{P} + (1-\mathrm{P})log_2((1-\mathrm{P})/(\mathrm{N}-1))\} \times (60/\mathrm{T}) \quad (1)$$

where N is the number of instruction sets, P is the classification accuracy and T is the consuming time for each selection. In offline experiments, the consuming time was 1.7s, 2.7s, 3.7s, 4.7s, and 5.7s for 1 to 5 rounds, respectively. The extra 4s of robot communication and movement time were required in the online experiments.

III. Results and Discussion

A. Offline BCI Performance

This section reported the accuracies and ITRs of the offline experiments. A leave-one-out cross-validation, which meant the data of 1 target was used as the test set and the data of the other 107 targets were used as the training set in each of the 108 validation steps, was adopted to ensure the robustness of the classification accuracy. For Fig. 5, it's obvious that the overall target accuracies depended on the accuracies of both sub-speller classification and pixel classification within a sub-speller. The pixel within a sub-speller has the lowest accuracy in Fig. 5a. However, the ITRs of the pixel classification within a sub-speller were higher than the other two ITRs in Fig. 5b. The reason for this is the large increase in the number of commands has led to an increase in the ITRs. As shown in Fig. 5, the average accuracy of target recognition had a steadily rising trend against the number of rounds, which increased from 87.04 % at 1 round to 96.66 % at 5 rounds. Fig. 5b shows the corresponding simulated online ITRs, which include another 0.7s of the cue time for each selection (but not including 4s of robot communication and movement time). The results showed the average ITR reached a maximum of 189.27 bits/min at 1 round. Therefore, only one round was used for the following online test.

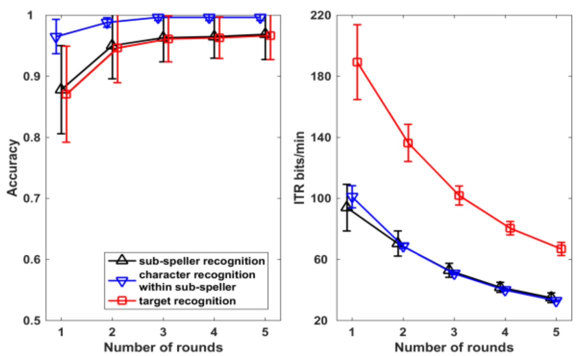

Figure 5. The accuracies and the corresponding ITRs of sub-speller recognition, pixel recognition within sub-spellers, and overall target recognition across subjects are plotted against the number of rounds, which are achieved by using the ensemble TRCA and hybrid EEG features. The error bars indicated standard errors.

B. Online BCI Performance

Table 1 lists the results of the online tests for the five subjects. All subjects were asked to write a Chinese character "福". The robot control system simply reproduced the results shown on the stimulation interface in the online experiments, so the target cue time was kept consistent with an offline cue time of 0.7s. For the writing of each stroke of Chinese character, a total of 5.7s is required (cue time: 0.7s; flicker time: 1s; robot communication and movement time: 4s). Subject 1 achieved the highest ITR of 71.10 bits/min with an accuracy of 100%. Four subjects achieved higher than 80% accuracy and higher than 50 bits/min of ITRs. The average accuracy was 87.23% and the average ITR was 56.85 bits/min across all subjects. These results demonstrated the feasibility and effectiveness of the proposed BCI-controlled robot system with a large instruction set and high pixel resolution.

TABLE 1. Results of Online Experiments

Subject	Consuming Time (s)	Selections (Correct /Total)	ACC. (%)	ITR (bits/min)
S1	5.7 (0.7+1+4)	32/32	100	71.10
S2	5.7 (0.7+1+4)	32/37	86.49	55.50
S3	5.7 (0.7+1+4)	32/35	91.43	60.58
S4	5.7 (0.7+1+4)	32/39	82.05	51.22
S5	5.7 (0.7+1+4)	32/42	76.19	45.87
Max	----	----	100	71.10
Min	----	----	76.19	45.87
Mean±STD	----	----	87.23±0.09	56.85±9.63

IV. CONCLUSION

This study developed a BCI-controlled robot system to "write" instead of "spell" Chinese characters. This work proposed for the first time a pixel-based writing method with the largest BCI instruction set for writing Chinese characters using a hybrid BCI. The average accuracy of the system reached 87.23% in the online tests, corresponding to an average ITR of 56.85 bits/min. The study results demonstrated that the proposed BCI system is a promising approach for writing ideogram (e.g. Chinese characters) and phonogram (e.g. English letter), which might lead to widespread real-world applications.

REFERENCES

[1] S. Gao et al., "Visual and auditory brain–computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 61, pp. 1436-1447, May. 2014.

[2] M. Xu et al., "A brain–computer interface based on miniature-event-related potentials induced by very small lateral visual stimuli," *IEEE Trans. Biomed. Eng.,* vol. 65, no. 5, pp. 1166-1175, 2018.

[3] K. Wang et al., "Enhance decoding of pre-movement EEG patterns for brain–computer interfaces," *Journal of neural engineering*, vol. 17, pp. 016033, 2020.

[4] E. Dong, et al. "Classification of multi-class motor imagery with a novel hierarchical SVM algorithm for brain–computer interfaces." *Medical & biological engineering & computing*,vol. 55, no.10, pp. 1809-1818, 2017.

[5] L. Xu, et al., "Cross-Dataset Variability Problem in EEG Decoding With Deep Learning." *Front. Hum. Neurosci.*, 2020.

[6] X. Xiao et al. "Discriminative canonical pattern matching for single-trial classification of ERP components." *IEEE Transactions on Biomedical Engineering*, 2019.

[7] M. Xu et al., "A hybrid BCI speller paradigm combining P300 potential and the SSVEP blocking feature," *J. Neural Eng.,* vol. 10, pp. 026001, Apr. 2013

[8] G. Townsend and V. Platsko, "Pushing the P300-based brain–computer interface beyond 100bpm: extending performance guided constraints into the temporal domain," *J. Neural Eng.,* vol. 13, pp. 026024, Apr. 2016.

[9] M. Nakanishi et al., "Enhancing Detection of SSVEPs for a High-Speed Brain Speller Using Task-Related Component Analysis," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 1, pp. 104-112, Jan, 2018.

[10] J.W. Minett et al., "An Assistive Communication Brain-Computer Interface for Chinese Text Input", *4th Int. Conf. on Bioinform. and Biomed. Eng.*, pp.1-4, June, 2010

[11] J. W. Minett et al., "A Chinese Text Input Brain–Computer Interface Based on the P300 Speller," *Int. J. Hum-Comput. Int.*, vol. 28, no. 7, pp. 472-483, 2012.

[12] Y. Yu et al., "A P300-Based Brain–Computer Interface for Chinese Character Input," *Int. J. Hum-Comput. Int.,* vol. 32, no. 11, pp. 878-884, 2016.

[13] Y. Wang et al., "A Benchmark Dataset for SSVEP-Based Brain-Computer Interfaces," *IEEE Trans Neural Syst. Rehabil. Eng.*, vol. 25, no. 10, pp. 1746-1752, Oct, 2017.

[14] M. Xu et. al, "Implementing over 100 command codes for a high-speed hybrid brain-computer interface using concurrent P300 and SSVEP features," *IEEE Trans. Biomed. Eng.,* 2020.

[15] J.R. Wolpaw et al., "Brain-computer interface technology: a review of the first international meeting". *IEEE Trans. Rehabil. Eng.*, vol. 8, pp. 164-173, Jun, 2000.