A Spectrally-Dense Encoding Method for Designing a High-Speed SSVEP-BCI With 120 Stimuli

Xiaogang Chen, *Member, IEEE*, Bingchuan Liu, *Student Member, IEEE*, Yijun Wang, *Member, IEEE*, and Xiaorong Gao, *Member, IEEE*

***Abstract*— The practical functionality of a brain-computer interface (BCI) is critically affected by the number of stimuli, especially for steady-state visual evoked potential based BCI (SSVEP-BCI), which shows promise for the implementation of a multi-target system for real-world applications. Joint frequency-phase modulation (JFPM) is an effective and widely used method in modulating SSVEPs. However, the ability of JFPM to implement an SSVEP-BCI system with a large number of stimuli, e.g., over 100 stimuli, remains unclear. To address this issue, a spectrally-dense JPFM (sJFPM) method is proposed to encode a broad array of stimuli, which modulates the low- and medium-frequency SSVEPs with a frequency interval of 0.1 Hz and triples the number of stimuli in conventional SSVEP-BCI to 120. To validate the effectiveness of the proposed 120-target BCI system, an offline experiment and a subsequent online experiment testing 18 healthy subjects in total were conducted. The offline experiment verified the feasibility of using sJFPM in designing an SSVEP-BCI system with 120 stimuli. Furthermore, the online experiment demonstrated that the proposed system achieved an average performance of 92.47 ± 1.83% in online accuracy and 213.23 ± 6.60 bits/min in online information transfer rate (ITR), where more than 75% of the subjects attained the accuracy above 90% and the ITR above 200 bits/min. This present study demonstrates the effectiveness of sJFPM in elevating the number of stimuli to more than 100 and extends our understanding of encoding a large number of stimuli by means of finer frequency division.**

***Index Terms*— Brain-computer interface (BCI), steady-state visual evoked potential (SSVEP), electroencephalography (EEG), joint frequency-phase modulation (JFPM), frequency division multiple access (FDMA), large number of stimuli.**

I. Introduction

A BRAIN-COMPUTER interface (BCI) offers a direct communication path between the brain and the outside world by translating the brain measurements associated with sensation, perception and cognition into commands or objective reports [1]. The BCI technology can be broadly categorized into invasive and non-invasive paradigms; invasive BCI is emerging in clinical applications and non-invasive BCI expands the scope to non-clinical daily applications. Among the non-invasive paradigms, steady-state visual evoked potential based BCI (SSVEP-BCI) [2], [3] is widely used in research along with its counterparts of P300-based BCI and motor imagery BCI. Compared with its counterparts, the SSVEP-BCI usually has a lower BCI-illiterate rate [4] and a higher information transfer rate (ITR) [3], which are attributed to the high signal-to-noise ratio (SNR) of SSVEP. Physiologically, SSVEP is a time-locked and frequency-tagged brain response elicited by flickers or checkerboards alternating at a certain stimulus frequency. The frequency-tagged attribute of SSVEP makes it a prime candidate for channel encoding, where the stimulus of each target can be efficiently encoded by the widely used joint frequency-phase modulation (JFPM) [3], [5], [6], [7], [8], [9], [10], [11].

The encoding approach of JFPM is critical in implementing a high-speed SSVEP-BCI and has significant implications in the development of the technology as well. Inspired by frequency-division multiple access (FDMA) in the communication system [12], JFPM configures each stimulus with a stimulation frequency that is equally spaced within a frequency interval. To further increase the separation ability between stimuli, an initial phase is added in the modulation and usually adjacent stimuli have distinct phase information. This joint modulation offers a high discriminability between stimuli for a short data length, which is a major advantage

Manuscript received 4 May 2022; revised 28 August 2022; accepted 19 September 2022. Date of publication 22 September 2022; date of current version 30 September 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 62171473, in part by the Strategic Priority Research Program of Chinese Academy of Science under Grant XDB32040200, in part by the Key Research and Development Program of Ningxia under Grant 2022CMG02026, in part by the Doctoral Brain+X Seed Grant Program of Tsinghua University, and in part by the Tianjin Municipal Science and Technology Plan Project under Grant 21JCYBJC01500. *(Xiaogang Chen and Bingchuan Liu contributed equally to this work.) (Corresponding authors: Xiaogang Chen; Yijun Wang.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Institutional Review Board of Tsinghua University under Application No. 20200020.

Xiaogang Chen is with the Institute of Biomedical Engineering, Chinese Academy of Medical Sciences and Peking Union Medical College, Tianjin 300192, China (e-mail: chenxg@bme.cams.cn).

Bingchuan Liu and Xiaorong Gao are with the Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing 100084, China (e-mail: lbc14@tsinghua.org.cn; gxr-dea@mail.tsinghua.edu.cn).

Yijun Wang is with the State Key Laboratory on Integrated Optoelectronics, Institute of Semiconductors, Chinese Academy of Sciences, Beijing 100083, China (e-mail: wangyj@semi.ac.cn).

This article has supplementary downloadable material available at https://doi.org/10.1109/TNSRE.2022.3208717, provided by the authors.

Digital Object Identifier 10.1109/TNSRE.2022.3208717

and critically important to implementing a high-performance BCI. Specifically, Chen *et al.* [3] reported a 40-target speller based on JFPM with an online ITR of up to 5.32 bits per second. Nakanishi *et al.* [6] developed an SSVEP-BCI system with 0.3-s visual stimuli and achieved an average online ITR of 325.33 ± 38.17 bits/min. Jiang *et al.* [8] incorporated a dynamic stopping strategy into the BCI system and attained an average ITR of 353.3 ± 67.1 bits/min with a peak of 460 bits/min. The evolution of the BCI systems also enhances our understanding of frequency recognition methods in SSVEP-BCI. For the frequency recognition, continuous efforts have been focused to improve the classification performance, such as extended CCA [13], task-related component analysis (TRCA) [6], multi-stimulus task-related component analysis (msTRCA) [14], and task-discriminant component analysis (TDCA) [15]. These efforts improved the ITR of the system by enhancing the classification accuracy and reducing the required selection time, both of which are important in the calculation of ITR, apart from the number of stimuli.

The practical functionality of the BCI system is associated with the number of stimuli. Therefore, compared with the efforts on target recognition, the studies on the number of stimuli merit special attention. For instance, a BCI system with two stimuli can be used for Yes/No answer selection [16]. A two-dimensional navigation or wheelchair control [17], [18] could be implemented with a small number of stimuli between four and nine. A moderate number of stimuli, i.e., 13 and 40 could be utilized to implement a BCI speller for digit [2] and character input [3], respectively. Thus, the elevated number of stimuli enables the BCI system to satisfy the need for more intricate tasks in real-world applications. Recently, research has focused on increasing the number of stimuli beyond one hundred in the BCI system to further augment the functionality. For example, Xu *et al.* [10] designed a hybrid BCI system with 108 stimuli using concurrent P300 and SSVEP features. By integrating 12 SSVEP stimuli with 3×3 P300 sub-spellers, 108 instruction sets were successfully encoded and were able to correctly select a target in 1.7 s. Chen *e al.* [19] devised an SSVEP-BCI system with 160 stimuli based on multiple frequencies sequential coding (MFSC) [20] and implemented a calibration-free system, by analyzing the combinations of the stimulus frequencies from target recognition. Very recently, Sun *e al.* [21] developed a code-modulated visual evoked potentials c-VEP based BCI system with 120 stimuli. Using four 31-bit pseudorandom codes, the proposed c-VEP based BCI system outperformed other c-VEP based systems in the number of stimuli and BCI performance. Although impressive progress has been achieved, this area of research remains in its infancy, with many issues that await further investigation. For instance, the average performances reported in prior works cannot simultaneously achieve high accuracy and high ITR. Additionally, it remains poorly understood whether the widely used JFPM is capable of implementing an SSVEP-BCI system with over 100 stimuli.

To address these issues, this study utilized JFPM to encode a large number of stimuli and validate a 120-target SSVEP-BCI system. Since SSVEP has a high SNR and the stimulus frequency of SSVEP has a wide frequency band [22], [23], it is theoretically sound to encode such a high volume of stimuli in the SSVEP spectrum. Under this assumption, a spectrally-dense JFPM (sJFPM) is hereby proposed by efficiently tagging a low- and medium-frequency band with a frequency interval of 0.1 Hz. A state-of-the-art task-related component analysis (TRCA) was then adopted in the target recognition. To the knowledge of the authors, this is the first study that expands the number of stimuli encoded by JFPM to over one hundred, which is considered a challenging problem by the previous study [19]. To validate the proposed system, an offline experiment was designed at first to verify the effectiveness of the proposed system and optimize the system parameters. In a further attempt to identify the ground-truth performance of the system, an online experiment testing 13 healthy subjects was then performed.

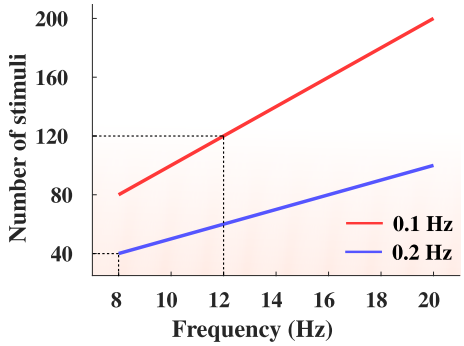

Fig. 1. The number of stimuli as a function of frequency in JFPM. The frequency denotes the lower limit of stimulus frequency. A frequency interval of 0.2 Hz (blue) and 0.1 Hz (red) are plotted. The two dashed lines denote the parameters of conventional 40-target SSVEP (40 stimuli) and the present study (120 stimuli), respectively.

II. Materials and Methods

A. Subjects

This study recruited 18 graduate students as healthy volunteers (eight males and ten females). The age of the subjects ranged from 23 to 28 with an average of 23.9 ± 1.6 years (mean ± standard deviation). Twelve of them participated in the offline experiment and 13 participated in the online experiment. Seven subjects participated in both experiments. All subjects were right-handed and had normal or corrected to normal vision. This study was approved by the institutional review board of Tsinghua University (NO. 20200020), and informed consent was signed by subjects before experimentation.

B. Spectrally-Dense Joint Frequency-Phase Modulation

An SSVEP-BCI brain speller was designed in this study with 120 stimuli, which were aligned in a 6×20 matrix. Based on JFPM, the frequency and initial phase information was encoded as follows:

$$\begin{aligned} f_{i,j} &= f_0 + [6(j-1) + i - 1] \cdot \Delta f \\ \Phi_{i,j} &= \Phi_0 + [6(j-1) + i - 1] \cdot \Delta\Phi \end{aligned} \tag{1}$$

where i (j) is the row (column) index of the stimuli, and Δf ($\Delta\Phi$) is the frequency (initial phase) interval that starts with the lower limit f_0 (Φ_0).

In conventional JFPM, the lower limit f_0 is configured to approximate the range of the stimulus frequencies f_r ($f_r = f_0 - \Delta f$), which prevents the spectral overlap between the fundamental and harmonic frequencies. For instance, a widely used configuration, where $f_0 = 8$ Hz and $\Delta f = 0.2$ Hz, is forming a 40-target SSVEP-BCI speller [3], [6], [24]. The lower limit f_0 can be increased and the frequency interval Δf can be decreased in JFPM to expand the number of stimuli Figure 1. To ensure a large number of encoded stimuli, a spectrally-dense JFPM (sJFPM) configuration was employed in this study by setting Δf and f_0 to 0.1 Hz and 12 Hz, respectively. In this fashion, the spectral band encoded for the stimulus consisted of 120 stimuli, ranging from 12 Hz to 23.9 Hz, The phase information was set as $\Phi_0 = 0, \Delta\Phi = 0.35\pi$ [3]. The details of the sJFPM configuration are summarized in Figure 2(**A**).

Figure 2(**B**) illustrates the user interface of the brain speller which was presented on the screen of a 48.9-inch LCD monitor (SAMSUNG; refresh rate: 120 Hz; resolution: 3840 × 1080 pixels). Each stimulus or target in the speller had a dimension of 173 × 129 pixels with a digit character from 1 to 120 at its center. The spacings between two adjacent stimuli were vertically 15 pixels and horizontally 19 pixels. The topmost black rectangle was set for result feedback.

A sampled sinusoidal stimulation method [25] was used to implement the sJFPM. In a stimulus sequence, the grayscale value of the i-th frame for the stimulus frequency f can be obtained by

$$s(f, \phi, i) = \frac{1}{2}\{1 + \sin[2\pi f(i/f_m) + \phi]\} \tag{2}$$

where f_m is the refresh rate of the monitor, i.e., 120 Hz in the study. Thus, a visual flicker was generated with a grayscale between 0 (dark) and 1 (highest luminance) in a sinusoidal manner. The visual stimuli were presented by Psychophysics Toolbox [26] in MATLAB (MathWorks, Inc.).

C. Offline Experiment

Six blocks of offline experiments were designed for system optimization. Each block comprised 120 trials in which the order was randomized, and there was one trial for each stimulus. Each trial lasted 4 s, including 1-s cue, 2-s visual stimulation and 1-s rest. In particular, trials started with a red square on top of a stimulus prompting for 1 s, and subjects were instructed to direct their attention to the prompted stimulus. Then all stimuli began to flicker simultaneously. Subjects were asked to look at the center of the prompted stimulus and avoid movement as well as eye blinking during the 2-s flickering. Finally, the speller paused for 1 s and subjects rested briefly. To avoid visual fatigue, there was a break of 5 min between two consecutive blocks. EEG data from the offline experiment were used for parameter selection to optimize the BCI system.

D. Online Experiment

An online experiment was conducted on a separate day after the offline experiment to validate the performance of the BCI system. Different from the offline experiment, in the online experiment the duration of visual stimulation was 0.7 s and gaze shift time for each stimulus was 1 s. In other words, the BCI system output a command in 1.7 s. The online experiment comprised nine blocks, including seven blocks for a training session and two blocks for a test session. No result feedback was provided in the training session as in the offline experiment. However, in the test session, an auditory feedback of a short beep was provided to the subject at the end of flickering if the SSVEP was correctly recognized by the target recognition algorithm. SSVEPs from the training session were used to train a model for target recognition in the test session.

E. Data Acquisition

This study recorded nine channels of EEG data using SynAmps2 (Neuroscan Inc., Charlotte, USA) for both offline and online experiments. The nine channels were from the classical occipital montage [24] in the international 10-20 system, i.e., Pz, POz/Oz, PO3/4, PO5/6 and O1/O2, which were also used for online analysis in target recognition. The impedances of the channels were maintained below 20 kΩ and the reference channel was set at Cz. The sampling rate was set at 1000 Hz, and EEG data were synchronized to the event triggers of the visual stimuli via a parallel port. EEG data were acquired in an electromagnetic shielding room to reduce environmental noise, and the power-line interference was removed by a hardware notch filter. The data were then downsampled to 250 Hz for offline and online analysis.

F. Data Analysis

This study used a state-of-the-art task-related component analysis (TRCA) [6] for target recognition. The performance of the proposed BCI system in target recognition was evaluated using classification accuracy and information transfer rate (ITR). The metric of ITR in bits per min (bits/min) is defined as [27]:

$$\begin{aligned} ITR = 60/T \cdot ([&\log_2 M + P\log_2 P \\ &+ (1-P)\log_2[(1-P)/(M-1)]) \end{aligned} \tag{3}$$

where M is the number of stimuli, P the classification accuracy, and T (s) the overall time for target selection and gaze shift. A gaze shift time of 1 s was used to calculate the ITR in both offline and online analyses.

1) Offline Analysis: First, parameter selection was conducted to determine the optimal number of sub-bands (N_{sb}) and stimulation duration (T_s). Based on the sJFPM scheme, the filter bank was designed with sub-bands ranging from $m \times 12$ Hz to 90 Hz, where m is the index of sub-band that ranged from 1 to N_{sb}. The stimulation duration T_s can be determined by a sliding window, in which the onset was set at $t_s + d$, where t_s is the starting time point of the stimulation. Note that d is the latency including the visual delay and system delay and it was set to 140 ms [3]. A parameter selection was performed for N_{sb} and T_s that varied between 1 and 7 and ranged from 0.2 s to 2 s with a step of 0.1 s, respectively. The parameters that yielded the best BCI performance were used in the online experiment.

Fig. 2. Virtual keyboard for the 120-target SSVEP-BCI speller. **(A)** Frequency (blue) and initial phase (red) encoded by spectrally-dense joint frequency-phase modulation (sJFPM) for each stimulus. **(B)** User interface of the speller comprising 120 stimuli, aligned in a 6×20 matrix.

The next step involved the investigation of BCI system performance under various calibration settings. Two settings were considered, i.e., different montage configuration and insufficient training data. A varying number of channels (N_{ch}) was selected for the montage configuration sequentially from the list of {Oz, O1, O2, POz, PO3, PO4, Pz, PO5, PO6}, where $N_{ch} = 3 - 9$ and the corresponding montages are detailed in the Supplementary Material (Table SI). For the insufficient training data, the number of training blocks (N_b) varied from 1 to 5. Under these settings, the accuracy at a data length of 2 s and the maximum average ITR across data lengths were used as the performance metrics.

To investigate the misclassification patterns in the target recognition, the output of the TRCA classifier, its intermediate features and the SNR of EEG trials were analyzed. The EEG trials with a 2 s data length were used for analysis. The confusion matrix was computed and the accuracy with respect to each stimulus was thereby obtained for all subjects. The R-squared statistic was used for intermediate features since it can well characterize the discrimination ability of stimuli in the feature space well [6], [28]. The R-squared feature was calculated from the correlation coefficients corresponding to the target class and nontarget classes. The SNR was calculated for each stimulus frequency using the following formula [9]:

$$SNR = 10\log_{10}\frac{\sum_{k=1}^{k=N_h} P(k \cdot f_n)}{\sum_{f=0}^{f=f_s/2} P(f) - \sum_{k=1}^{k=N_h} P(k \cdot f_n)} \quad (4)$$

where $P(f)$ is the spectral power for the frequency f, and f_n the stimulus frequency and N_h the number of harmonics, which is usually set to 5 [9].

Furthermore, the effect of the frequency interval on classification performance was explored. The stimulus frequencies were sorted in ascending order and the associated stimuli were sampled by an interval of δ Hz, where $\delta = 0.1, 0.2, 0.3, 0.4, 0.6, 0.8, 1$ to simulate a variety of frequency intervals assigned by JFPM. Specifically for each frequency interval, there were $10 \cdot \delta$ subsets that comprised N_c stimuli, and we performed a N_c-target classification was performed on each subject, where $N_c = 12 \cdot \delta^{-1}$. The performance of classification was then evaluated by the maximum average ITR across subsets, blocks and subjects.

A 6-fold leave-one-block-out cross validation was performed in the offline analysis, where in each fold one block was used as the test data and the remaining were used as training data. For multiple comparisons in the offline analysis, a repeated measures analysis of variance (ANOVA) was applied. To account for the violation of sphericity, as assessed by Mauchly's test of sphericity, Greenhouse-Geisser correction was then employed. Post hoc comparisons using t-test with Bonferroni correction were conducted when there was a statistically significant main effect ($p < .05$). The statistical analyses were performed in SPSS Statistics 26 (IBM, Armonk, NY, USA). Unless otherwise stated, data were presented as mean $\pm$ standard error in this study.

2) Online Analysis: The number of sub-band N_{sb} in the online TRCA model was determined by the parameters chosen in the offline analysis, where $N_{sb} = 4$. The model detected online EEG data with a data length of 0.7 s after latency correction ($d = 140$ ms) were recognized by the model to provide real-time result feedback. The online classification accuracy and ITR were then computed for each subject.

III. Results

A. Offline Result

The average classification performance (**A**: accuracy, **B**: ITR) for a varying N_{sb} and T_s in the parameter selection are

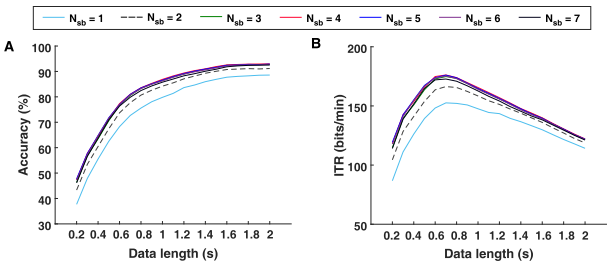

Fig. 3. Change in classification performance with respect to two parameters, i.e., the number of sub-band N_{sb} and data length T_s in parameter selection. (**A**) Average accuracy. (**B**) Average ITR. Data lengths from 0.2 s to 2 s with an interval of 0.1 s were used for the evaluation.

illustrated in Figure 3. Visual inspection reveals a noticeable difference in the classification performance can be observed as the parameter changes. As assessed by two-way (sub-band × data length) repeated measures ANOVA, there was a statistically significant interaction between the sub-band and data length on accuracy, $F(3.182, 35.001) = 7.322$, $p = .001$, and on ITR, $F(3.356, 36.912) = 15.078$, $p = .003$. An optimal $T_s = 0.7$ that achieved the highest ITR across data lengths could be determined from the result of ITR in Figure 3 **B**. To pinpoint the optimal number of sub-bands N_{sb}, T_s was fixed and the change in accuracy was delineated, as shown in Figure 4 **A**. One-way repeated measures ANOVA revealed that there was a significant difference in accuracies between sub-bands, $F(1.119, 12.306) = 13.371$, $p = .003$. A sub-band parameter of $N_{sb} = 4$ yielded the highest accuracy, i.e., $81.02 \pm 4.01\%$. Using the optimal parameter, the average accuracy and ITR across data lengths together with individual results are further illustrated in Figure 4 **B** and Figure 4 **C**, respectively. Here, each dashed line represents a subject and the blue line is their average. The result showed that the BCI performance exhibited an inter-subject variation. For instance, the subject with excellent performance (S5) achieved a maximum ITR of 253.53 bits/min, and the subject with poor performance (S6) achieved a maximum ITR of 132.68 bits/min. As for the average result, the average accuracy at 2 s was 92.99 ± 1.99 %, and the maximum average ITR was 176.17 ± 12.99 bits/min at 0.7 s.

The effect of various occipital montages on the BCI performance is illustrated in Figure 5 (**A**: accuracy, **B**: ITR). Here, each dashed line represents a subject to show their individual performance, and the solid line is their average. One-way repeated measures ANOVA revealed a statistically significant difference between montages in accuracy, $F(2.145, 23.592) = 24.809$, $p < .001$, and in ITR, $F(2.583, 28.411) = 52.815$, $p < .001$. The BCI performance improved significantly by increasing the number of channels in montages. The usage of three channels (Oz, O1 and O2) achieved an accuracy of 97.08% for subjects with excellent BCI performance (S5), which was above the average (96.25%) using nine channels. The addition of a channel (e.g., Pz) decreased the performance for some subjects, suggesting the individual difference in the topography of SSVEP responses.

Figure 6 illustrates the change in performance with a varying number of training blocks. One-way repeated measures ANOVA revealed a statistically significant difference between different number of blocks in accuracy, $F(1.34, 14.743) = 43.127$, $p < .001$, and in ITR, $F(1.714, 18.853) = 118.439$, $p < .001$. The result showed that the BCI performance increased significantly as more training data were available. By leveraging two blocks of training data, the average ITR surpassed 100 bits/min, i.e., 141.48 ± 16.07 bits/min. Using three blocks of training data, the average accuracy surpassed 80%, i.e., $83.67 \pm 4.13\%$. Of note, with insufficient training data of one block, three subjects achieved relatively higher performance, i.e., S5 (97.64%), S8 (78.75%) and S7 (77.08%).

To investigate the misclassification pattern, the confusion matrix was constructed as shown in Figure 7 **A**. The number of trials greater than five, which were 72 in total, were colored blue for contrast. A diagonal due to the high accuracy at 2 s is clear, and errors in adjacent stimuli could also be identified. On closer examinations of errors, Figure 7 **B** and **C** depict the accuracy and R-squared maps for each stimulus, respectively. Here, the accuracies and R-squared features were arranged according to the user interface of the speller. For the accuracy map, stimuli in the center left region of the speller were recognized with higher accuracy (>90%), whereas the right part had a higher error rate. A similar pattern was observed in the R-squared map, where a higher value as an indicator of better discrimination ability was detected in the left center region. The distribution of the SNR was further delineated to probe into the causal role. Figure 7 **D** illustrates the bar plot of the SNR values with respect to each stimulus frequency, indicating a decreasing tendency ($r = -0.964$, $p < .001$) in SNR as the stimulus frequency increases. To reveal its spatial distribution, the SNR map was illustrated in a similar fashion in Figure 7 **E**. The result showed that the SSVEP trials evoked by the stimuli in the left region of the speller (column $1-10$) have a significantly higher SNR than that in the right region (left: -10.433 ± 0.138 dB; right: -13.109 ± 0.127 dB; $p < .001$).

The relationship between frequency interval and simulated BCI performance is illustrated in Figure 8. One-way repeated measures ANOVA revealed a statistically significant difference in ITR between frequency intervals, $F(1.445, 15.890) = 80.303$, $p < .001$. Post hoc pairwise comparisons with Bonferroni correction revealed that there was a significant difference between $\delta = 0.1$ and $\delta = 0.3, 0.4, 0.6, 0.8, 1$, with all $p < .05$. However, no significant difference was found between $\delta = 0.1$ (ITR: 176.53 ± 12.97 bits/min) and $\delta = 0.2$ (ITR: 177.05 ± 12.40 bits/min), with $p = 1$, Bonferroni corrected.

B. Online Result

Table I lists the individual and average classification accuracy and ITR recorded in the online experiment among 13 subjects. The average accuracy achieved was $92.47 \pm 1.83\%$, while the ITR for the BCI system was 213.23 ± 6.60 bits/min. As for the individual performance, over 75% of the subjects (10/13) achieved an online accuracy above 90% and an online

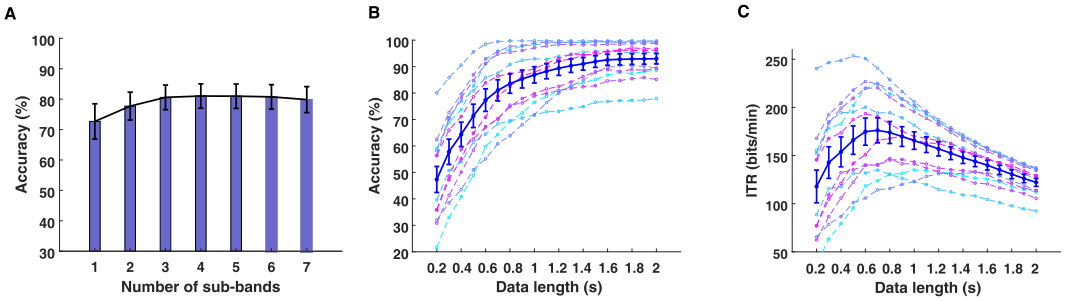

Fig. 4. Change in classification performance with respect to one of the parameters, i.e., the number of sub-band N_{sb} (**A**) and data length T_s (**B**, **C**) in parameter selection. (**A**) Average accuracy with a varying N_{sb} when the optimal parameter of T_s is fixed at $T_s = 0.7$. (**B**) Average accuracy with a varying T_s when the optimal parameter of N_{sb} is fixed at $N_{sb} = 4$. (**C**) Average ITR with a varying T_s when the optimal parameter of N_{sb} is fixed at $N_{sb} = 4$. In (**B**) and (**C**), the thin lines represent the subjects and the thick line is their average. The shaded area represents the standard error. Data lengths from 0.2 s to 2 s with an interval of 0.1 s were used for the evaluation.

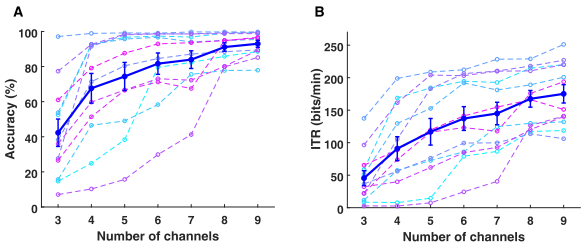

Fig. 5. Change in classification performance with a varying number of channels. (**A**) Classification accuracy at 2 s. (**B**) Maximum average ITR across data lengths. Each dashed line represents a subject and the solid line is their average. The shaded area represents the standard error.

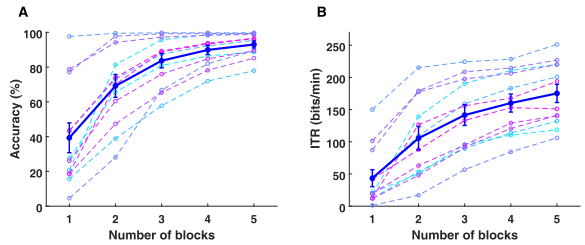

Fig. 6. Change in classification performance with a varying number of training blocks. (**A**) Classification accuracy at 2 s. (**B**) Maximum average ITR across data lengths. Each dashed line represents a subject and the solid line is their average. The shaded area represents the standard error.

ITR above 200 bits/min. Four subjects (S4, S6, S8 and S11) accomplished an online accuracy above 98% and an online ITR above 230 bits/min. The subject with the highest performance was S8, who attained an online accuracy of 98.75% and an online ITR of 237.31 bits/min. A demo video of the BCI system can be found at https://youtu.be/lAt2arWT9wo.

IV. Discussion

In this study, a spectrally-dense joint frequency-phase modulation (sJFPM) was utilized to encode 120 stimuli with a frequency interval of 0.1 Hz. First, a BCI system was designed by sJFPM, and the parameters were optimized in an offline experiment. Subsequently, an online experiment was performed on 13 subjects to validate the proposed system. The results demonstrated that the 120-target brain speller could achieve an average classification accuracy above 90%

TABLE I
Results of the Online Experiment Testing 13 Subjects. The Results for Subjects With the Top Four Online Performances Are Marked in Bold

Subject	Accuracy (%)	ITR (bits/min)
S1	94.58	219.87
S2	82.92	178.92
S3	90.42	204.37
S4	**98.33**	**235.40**
S5	96.25	226.50
S6	**98.33**	**235.40**
S7	77.92	163.15
S8	**98.75**	**237.31**
S9	92.92	213.51
S10	85.83	188.52
S11	**98.33**	**235.40**
S12	91.67	208.89
S13	95.83	224.81
Mean±sem	**92.47±1.83**	**213.23±6.60**

and an average ITR above 200 bits/min. The result of the proposed system substantially outperformed the existing BCI systems with over 100 stimuli in accuracy [10], [19], [21] and in ITR [10], [19]. The superiority of the proposed system in performance verifies the effectiveness of the sJFPM in designing a high-speed SSVEP-BCI with a large number of stimuli.

Three issues concerning sJFPM and target recognition are further discussed in the following paragraph. First, as for the frequency interval in sJFPM, the BCI performance was measured by means of sampling from the stimuli, suggesting that the traditional JFPM could be scaled to sJFPM to expand the number of stimuli without inhibiting the performance in ITR. The frequency interval of 0.2 Hz is widely used in the implementation of high-speed SSVEP-BCI spellers [3], [6], [9], [24], [29], whereas the frequency interval of 0.1 Hz has received little attention [30]. The usage of 0.1-Hz frequency interval enables the BCI system to configure a broad array of stimuli, and as suggested by the result in Figure 8, the ITR of the system does not suffer from a degradation. This phenomenon of ITR is attributable to the compromise between the number of stimuli M and classification accuracy P in Eq. (3), in which the greater number of stimuli would inevitably

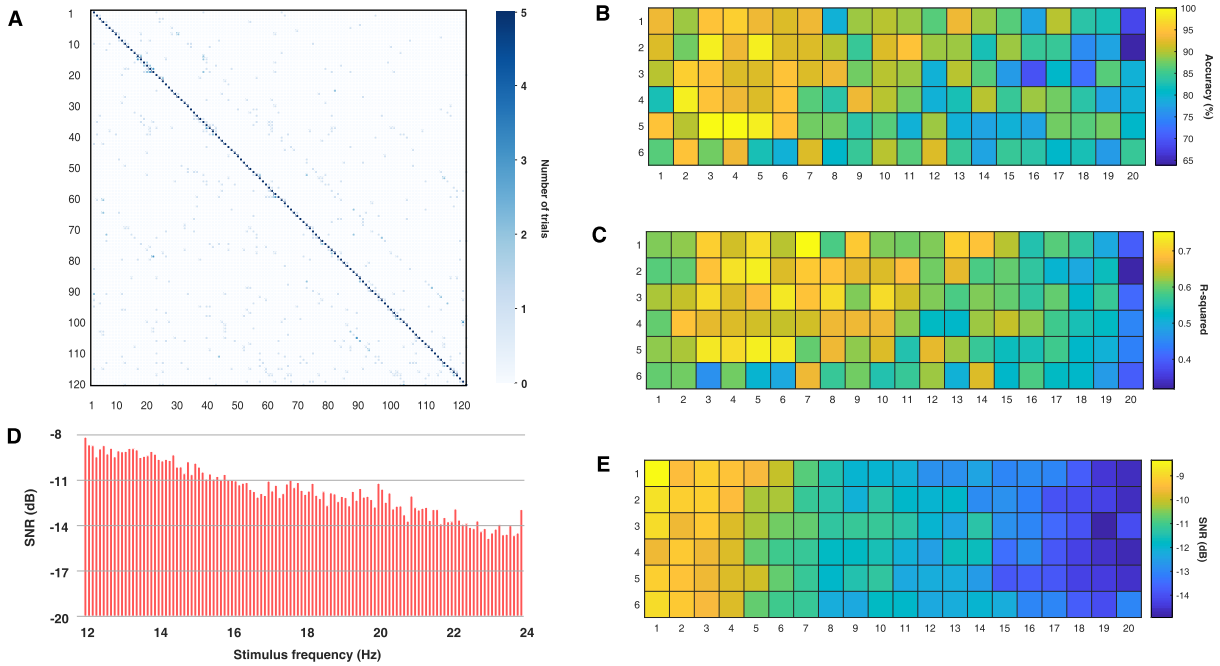

Fig. 7. Error analysis in classification. **(A)** Confusion matrix. There are 72 trials in total (12 subjects × 6 blocks) in this confusion matrix and the number of trials greater than five was colored blue for illustrative purposes. **(B)** The accuracy map for each stimulus. **(C)** The R-squared map for each stimulus. **(D)** Bar plot of the SNR values for each stimulus frequency. **(E)** The SNR map for each stimulus. EEG data with a data length of 2 s were used for analysis, and the values of accuracy, R-squared, and SNR were arranged according to the user interface of the speller.

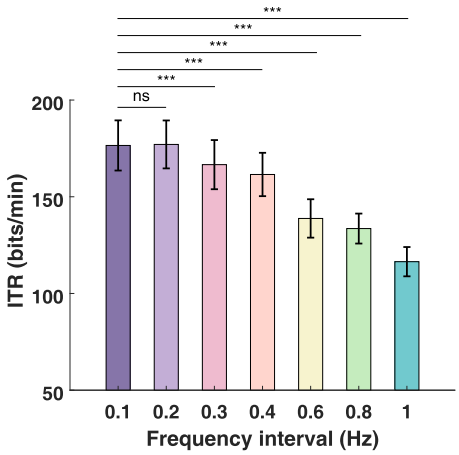

Fig. 8. Change in the maximum average ITR for various frequency intervals. The asterisks indicate a pairwise statistically significant difference (*$p < .05$, **$p < .01$, ***$p < .001$, Bonferroni corrected).

increase the error rate. For instance, the average accuracy corresponding to the maximum average ITR is $81.02 \pm 4.01\%$ at 0.7 s and $83.88 \pm 3.73\%$ at 0.5 s for 0.1-Hz and 0.2-Hz frequency intervals, respectively. Second, the error analysis in Figure 7 suggests that the higher error rate in the right region of the speller is caused by the lower discrimination ability for these stimuli, which is in part further due in part to the associated lower SNRs. Specifically, the R-squared statistics in Figure 7 accounted for 56.0% of the variation in the accuracy with adjusted $R^2 = .556$, $F(1, 118) = 150.294$, $p < .001$, and the SNRs accounted for 46.9% of the variation in the accuracy with adjusted $R^2 = .465$, $F(1, 118) = 104.257$, $p < .001$. The lower SNR value for the stimuli in the right region originated from the property of high-frequency SSVEP, in which SSVEPs with higher stimulus frequencies produce lower SNRs [22], [23] (Figure 7 **D**). Third, as for the target recognition, TRCA was used to validate the system validation, and other state-of-the-art methods [15] could be applied to further reduce the error rate.

Compared with the existing methods for a large number of stimuli in BCI, the present study based on the JFPM encoding is characterized by the following. In retrospect, the JFPM encoding method has long demonstrated its excellence in the implementation of a high-speed BCI. For the 40-target BCIs, previous studies [3], [6], [8] reported that BCI systems based on JFPM could achieve an average ITR of over 300 bits/min, which is the state-of-the-art performance in non-invasive BCIs to the knowledge of the authors. For the regime of a large number of stimuli, the effectiveness of JFPM was further demonstrated in the present study by implementing a high-speed BCI, which resolves the doubt about the feasibility of a high volume of stimulus frequencies for BCIs. The high-speed performance of JFPM is attributable to the fact that JFPM elicits a remarkably high SNR of brain signals (i.e., continuous SSVEP) compared with other methods in visual BCIs. In contrast to the P300 [10] and c-VEP [21] signals, SSVEP boasts a high single-trial SNR, which removes the hassle of repetition of visual stimuli multiple times and thereby increases the ITR of the system. In comparison with MFSC [19], JFPM provides a continuous encoding of SSVEP and elevates the SNR as the data length increases, whereas in MFSC the transition of stimulus frequencies interrupts the steady state of VEP

Fig. 9. A keyboard interface for one-keystroke Chinese character input. Here each stimulus denotes 1−4 pinyin-based syllables, each of which corresponds to Chinese characters for selection. The alphabet above the first line of stimuli denotes the initial letter of the syllable for each column.

and possibly offsets the amplitude of brain response. Because each stimulus in sJFPM is tagged by a stimulus frequency, a wide range of existing target recognition methods [6], [13], [14], [15], [24], [31] for SSVEP-BCI are applicable to the sJFPM-based system, while in MFSC the majority of the target recognition methods should be redesigned and custom adapted.

The combination of high-speed performance together with a large volume of stimuli make it possible for the BCI system to enable new applications. First, the large array of stimuli aligned densely in the plane would possibly expand the utility of BCI from selecting to locating. Since each target in the array corresponds to a spatial location or a coordinate, the coordinate information associated with the user intent could be decoded by the target recognition in SSVEP-BCI and then used for a wide range of applications, such as map location, spatial navigation, digital sandtable control. Second, the coordinates of the high-density array could be utilized to create the semantics of the stimuli, rather than character input from a single target. For instance, a hybrid BCI based on P300 and SSVEP was used to control a robotic arm with the coordinate of 108 stimuli and write a Chinese character by strokes [32]. Thus the BCI with a large number of stimuli can provide a means for people with motor disabilities to write, paint and create. Third, a syllable-based input method could be developed in the BCI system by leveraging a large number of stimuli, as an extension to the traditional brain speller for character input [3]. For example the Chinese language which has 416 syllables in total. With a large number of stimuli in BCIs, the syllables can be represented by the stimuli and users can enter a syllable in one keystroke. Figure 9 illustrates a keyboard interface tailored for Chinese character input and its designing procedure is provided in the Supplementary Material. Here each stimulus is mapped to 1−4 syllables and an alphabet is presented above the first line of stimuli to indicate the initial letter of the syllables. Users can find an intended stimulus by first referring to the alphabet and then selecting from only a few stimuli (no more than 7 stimuli). By taking advantage of the speller with 120 stimuli, the efficiency of Chinese character input could be significantly boosted.

The contribution of the present study is three-fold. First, methodologically, we extended the conventional JFPM to the spectrally-dense regime that prevents spectral overlap and encodes a large number of stimuli in the low- and medium-frequency band. In system implementation, we successfully developed an SSVEP-BCI system with 120 stimuli using sJFPM and the proposed system can achieve a high-speed performance as validated in online experiment. For practical applications, the proposed BCI system has the potential to enable a variety of new applications, e.g., BCI systems for one-keystroke Chinese character input. Nevertheless, this study is proposed as a proof-of-principle demonstration, which has its limitations, leaving room for its improvement in future work. To improve the practical utility of the BCI system, its ease of use can be improved in terms of EEG cap, standard monitor and the effort in reducing calibration time. Dry EEG cap [33] and pre-gelled EEG cap [34] are more practical and can be used to replace the gel-based EEG cap in real-world applications. A standard monitor can be employed for stimulus presentation in our future work to make the proposed BCI system more readily available. Cross-stimulus transfer learning [35] could be employed to overcome the calibration burden brought about by the increased number of stimuli. Other transfer learning approaches in SSVEP-BCI [36], [37], [38] have the potential to further increase the system performance by leveraging data from other subjects [36], [37] or other sources [38]. Apart from the effort in reducing calibration time, visual fatigue can be mitigated by designing a paradigm with more visual comfort [39]. For the improvement of BCI performance, more electrodes can be employed to further enhance the ITR of the 120-target system [15]. Furthermore, a dynamic stopping strategy [40] could be implemented to tackle the individual difference in parameter selection and thereby enhance the ITR of the system.

V. Conclusion

The present study proposed a spectrally-dense joint frequency-phase modulation (sJFPM) encoding method to design a high-speed steady-state visual evoked potential based brain-computer interface (SSVEP-BCI) system with 120 stimuli. Two experiments, an offline and an online, involving 18 subjects were conducted to optimize and validate the system performance. The results demonstrated that the proposed 120-target BCI system based on sJFPM could achieve an online accuracy of $92.47 \pm 1.83\%$ and an online ITR of 213.23 ± 6.60 bits/min. By means of using finer frequency division to encode a high volume of stimuli, the present study provides insight into the JFPM method and offers an opportunity for BCI to be involved in new BCI applications, contributing to the effort for developing novel non-invasive BCI systems.

References

[1] X. Gao *et al.*, "Interface, interaction, and intelligence in generalized brain–computer interfaces," *Trends Cogn. Sci.*, vol. 25, no. 8, pp. 671–684, 2021.

[2] M. Cheng, X. Gao, S. Gao, and D. Xu, "Design and implementation of a brain-computer interface with high transfer rates," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 10, pp. 1181–1186, Oct. 2002.

[3] X. Chen, Y. Wang, M. Nakanishi, X. Gao, T.-P. Jung, and S. Gao, "High-speed spelling with a noninvasive brain–computer interface," *Proc. Nat. Acad. Sci. USA*, vol. 112, no. 44, Nov. 2015, Art. no. 201508080.

[4] M.-H. Lee *et al.*, "EEG dataset and OpenBMI toolbox for three BCI paradigms: An investigation into BCI illiteracy," *GigaScience*, vol. 8, no. 5, Jan. 2019.

[5] B. Wittevrongel and M. M. Van Hulle, "Frequency- and phase encoded SSVEP using spatiotemporal beamforming," *PLoS ONE*, vol. 11, no. 8, Aug. 2016, Art. no. e0159988.

[6] M. Nakanishi, Y. Wang, X. Chen, Y. Wang, X. Gao, and T.-P. Jung, "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.

[7] B. Wittevrongel *et al.*, "Representation of steady-state visual evoked potentials elicited by luminance flicker in human occipital cortex: An electrocorticography study," *NeuroImage*, vol. 175, no. 15, pp. 315–326, Jul. 2018.

[8] J. Jiang, E. Yin, C. Wang, M. Xu, and D. Ming, "Incorporation of dynamic stopping strategy into the high-speed SSVEP-based BCIs," *J. Neural Eng.*, vol. 15, no. 4, Aug. 2018, Art. no. 046025.

[9] B. Liu, X. Huang, Y. Wang, X. Chen, and X. Gao, "BETA: A large benchmark database toward SSVEP-BCI application," *Frontiers Neurosci.*, vol. 14, p. 627, Jun. 2020.

[10] M. Xu, J. Han, Y. Wang, T.-P. Jung, and D. Ming, "Implementing over 100 command codes for a high-speed hybrid brain-computer interface using concurrent P300 and SSVEP features," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 11, pp. 3073–3082, Nov. 2020.

[11] B. Wittevrongel *et al.*, "Practical real-time MEG-based neural interfacing with optically pumped magnetometers," *BMC Biol.*, vol. 19, no. 1, pp. 1–15, Dec. 2021.

[12] S. Gao, Y. Wang, X. Gao, and B. Hong, "Visual and auditory brain–computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 5, pp. 1436–1447, May 2014.

[13] M. Nakanishi, Y. Wang, Y. Wang, Y. Mitsukura, and T. Jung, "A high-speed brain speller using steady-state visual evoked potentials," *Int. J. Neural Syst.*, vol. 24, no. 6, Sep. 2014, Art. no. 1450019.

[14] C. M. Wong *et al.*, "Learning across multi-stimulus enhances target recognition methods in SSVEP-based BCIs," *J. Neural Eng.*, vol. 17, no. 1, Jan. 2020, Art. no. 016026.

[15] B. Liu, X. Chen, N. Shi, Y. Wang, S. Gao, and X. Gao, "Improving the performance of individually calibrated SSVEP-BCI by task-discriminant component analysis," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1998–2007, 2021.

[16] X. Chen, N. Hu, Y. Wang, and X. Gao, "Validation of a brain-computer interface version of the digit symbol substitution test in healthy subjects," *Comput. Biol. Med.*, vol. 120, May 2020, Art. no. 103729.

[17] Y. Li, J. Pan, F. Wang, and Z. Yu, "A hybrid BCI system combining P300 and SSVEP and its application to wheelchair control," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 11, pp. 3156–3166, Nov. 2013.

[18] A. Cruz, G. Pires, A. Lopes, C. Carona, and U. J. Nunes, "A self-paced BCI with a collaborative controller for highly reliable wheelchair driving: Experimental tests with physically disabled individuals," *IEEE Trans. Human-Mach. Syst.*, vol. 51, no. 2, pp. 109–119, Apr. 2021.

[19] Y. Chen, C. Yang, X. Ye, X. Chen, Y. Wang, and X. Gao, "Implementing a calibration-free SSVEP-based BCI system with 160 targets," *J. Neural Eng.*, vol. 18, Jun. 2021, Art. no. 046094.

[20] Y. Zhang, P. Xu, T. Liu, J. Hu, R. Zhang, and D. Yao, "Multiple frequencies sequential coding for SSVEP-based brain-computer interface," *PLoS ONE*, vol. 7, no. 3, Mar. 2012, Art. no. e29519.

[21] Q. Sun, L. Zheng, W. Pei, X. Gao, and Y. Wang, "A 120-target brain-computer interface based on code-modulated visual evoked potentials," *J. Neurosci. Methods*, vol. 375, Jun. 2022, Art. no. 109597.

[22] D. Regan, "Evoked potentials and evoked magnetic fields in science and medicine," *Hum. Brain Electrophysiol.*, pp. 59–61, 1989.

[23] C. S. Herrmann, "Human EEG responses to 1–100 Hz flicker: Resonance phenomena in visual cortex and their potential correlation to cognitive phenomena," *Experim. Brain Res.*, vol. 137, nos. 3–4, pp. 346–353, 2001.

[24] X. Chen, Y. Wang, S. Gao, T.-P. Jung, and X. Gao, "Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain–computer interface," *J. Neural Eng.*, vol. 12, no. 4, Jun. 2015, Art. no. 046008.

[25] X. Chen, Z. Chen, S. Gao, and X. Gao, "A high-ITR SSVEP-based BCI speller," *Brain-Comput. Interfaces*, vol. 1, nos. 3–4, pp. 181–191, Sep. 2014.

[26] D. H. Brainard, "The psychophysics toolbox," *Spatial Vis.*, vol. 10, no. 4, pp. 433–436, 1997.

[27] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain–computer interfaces for communication and control," *Clin. Neurophysiol.*, vol. 113, no. 6, pp. 767–791, 2002.

[28] Y. Wang, M. Nakanishi, Y.-T. Wang, and T.-P. Jung, "Enhancing detection of steady-state visual evoked potentials using individual training data," in *Proc. 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2014, pp. 3037–3040.

[29] Y. Wang, X. Chen, X. Gao, and S. Gao, "A benchmark dataset for SSVEP-based brain-computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.* vol. 25, no. 10, pp. 1746–1752, Nov. 2017.

[30] H.-J. Hwang, J.-H. Lim, Y.-J. Jung, H. Choi, S. W. Lee, and C.-H. Im, "Development of an SSVEP-based BCI spelling system adopting a QWERTY-style LED keyboard," *J. Neurosci. Methods*, vol. 208, no. 1, pp. 59–65, Jun. 2012.

[31] G. Bin, X. Gao, Z. Yan, B. Hong, and S. Gao, "An online multi-channel SSVEP-based brain–computer interface using a canonical correlation analysis method," *J. Neural Eng.*, vol. 6, no. 4, Jun. 2009, Art. no. 046002.

[32] J. Han *et al.*, "'Write'but not 'spell'Chinese characters with a BCI-controlled robot," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 4741–4744.

[33] X. Xing *et al.*, "A high-speed SSVEP-based BCI using dry EEG electrodes," *Sci. Rep.*, vol. 8, no. 1, pp. 1–10, Dec. 2018.

[34] W. Pei *et al.*, "A pre-gelled EEG electrode and its application in SSVEP-based BCI," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 843–850, 2022.

[35] C. M. Wong *et al.*, "Transferring subject-specific knowledge across stimulus frequencies in SSVEP-based BCIs," *IEEE Trans. Autom. Sci. Eng.*, vol. 18, no. 2, pp. 552–563, Apr. 2021.

[36] C. M. Wong *et al.*, "Inter- and intra-subject transfer reduces calibration effort for high-speed SSVEP-based BCIs," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 10, pp. 2123–2135, Oct. 2020.

[37] K.-J. Chiang, C.-S. Wei, M. Nakanishi, and T.-P. Jung, "Boosting template-based SSVEP decoding by cross-domain transfer learning," *J. Neural Eng.*, vol. 18, no. 1, Feb. 2021, Art. no. 016002.

[38] B. Liu, X. Chen, X. Li, Y. Wang, X. Gao, and S. Gao, "Align and pool for EEG headset domain adaptation (ALPHA) to facilitate dry electrode based SSVEP-BCI," *IEEE Trans. Biomed. Eng.*, vol. 69, no. 2, pp. 795–806, Feb. 2022.

[39] X. Zhao, Z. Wang, M. Zhang, and H. Hu, "A comfortable steady state visual evoked potential stimulation paradigm using peripheral vision," *J. Neural Eng.*, vol. 18, no. 5, Oct. 2021, Art. no. 056021.

[40] E. Yin, Z. Zhou, J. Jiang, Y. Yu, and D. Hu, "A dynamically optimized SSVEP brain–computer interface (BCI) speller," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 6, pp. 1447–1456, Jun. 2015.