Journal of Neural Engineering

PAPER

A high-speed hybrid brain-computer interface with more than 200 targets

Jin Han[1] , Minpeng Xu[1,2,*], Xiaolin Xiao[1,2] , Weibo Yi[3], Tzyy-Ping Jung[1,4] and Dong Ming[1,2,*]

1 Department of Biomedical Engineering, College of Precision Instruments and Optoelectronics Engineering, Tianjin University, Tianjin 300072, People's Republic of China
2 Academy of Medical Engineering and Translational Medicine, Tianjin University, Tianjin 300072, People's Republic of China
3 Beijing Machine and Equipment Institute, Beijing 100854, People's Republic of China
4 Swartz Centre for Computational Neuroscience, University of California, San Diego, CA, United States of America
* Authors to whom any correspondence should be addressed.

E-mail: minpeng.xu@tju.edu.cn and richardming@tju.edu.cn

Abstract

Objective. Brain-computer interfaces (BCIs) have recently made significant strides in expanding their instruction set, which has attracted wide attention from researchers. The number of targets and commands is a key indicator of how well BCIs can decode the brain's intentions. No studies have reported a BCI system with over 200 targets. *Approach.* This study developed the first high-speed BCI system with up to 216 targets that were encoded by a combination of electroencephalography features, including P300, motion visual evoked potential (mVEP), and steady-state visual evoked potential (SSVEP). Specifically, the hybrid BCI paradigm used the time-frequency division multiple access strategy to elaborately tag targets with P300 and mVEP of different time windows, along with SSVEP of different frequencies. The hybrid features were then decoded by task-discriminant component analysis and linear discriminant analysis. Ten subjects participated in the offline and online cued-guided spelling experiments. Other ten subjects took part in online free-spelling experiments. *Main results.* The offline results showed that the mVEP and P300 components were prominent in the central, parietal, and occipital regions, while the most distinct SSVEP feature was in the occipital region. The online cued-guided spelling and free-spelling results showed that the proposed BCI system achieved an average accuracy of $85.37\% \pm 7.49\%$ and $86.00\% \pm 5.98\%$ for the 216-target classification, resulting in an average information transfer rate (ITR) of 302.83 ± 39.20 bits min^{-1} and 204.47 ± 37.56 bits min^{-1}, respectively. Notably, the peak ITR could reach up to 367.83 bits min^{-1}. *Significance.* This study developed the first high-speed BCI system with more than 200 targets, which holds promise for extending BCI's application scenarios.

1. Introduction

Brain-computer interfaces (BCIs) provide a special pathway for the brain to communicate directly with the external environment without relying on the peripheral nervous system [1–4]. Visual BCI (v-BCI) systems have recently made significant progress and become one of the most popular paradigms. Specifically, the most frequently used brain control signals for v-BCI are motion visual evoked potential (mVEP) [5, 6], P300 [7, 8], steady-state visual evoked potential (SSVEP) [9, 10], and their hybrid features [11, 12], which have shown great promise in practical applications.

Information transfer rate (ITR), a basic BCI index that expresses the BCI's communication ability, is determined by speed, accuracy, and number of targets [13, 14]. Since Vidal first proposed BCI in 1973 [15], researchers have initially focused on system accuracy to demonstrate the feasibility of translating brain signals to computer instructions. Specifically, in 1988, Farwell and Donchin designed a P300-speller that encoded 36 targets for the brain to output intents with an average accuracy of more than 85% [16]. In 1996,

Calhoun and McMillan developed an SSVEP-based BCI system with an accuracy rate of over 80% [17]. The speed of brain outputs later became a hot topic as researchers began to focus more on the effectiveness of BCIs. In 2015, the response time of the BCI system was shortened to 1 s [13]. In 2018, the speed of BCI outputs was further decreased to 0.8 s per target [18]. The ITR could only be marginally improved by addressing the BCI speed then, so researchers tried to increase the target number to increase the BCI's capacity for encoding brain intents. In 2016, Townsend and Platsko developed a 72-target BCI system that first pushed the P300-based BCIs past 100 bits min^{-1} [19]. In 2020, Xu *et al* developed the first high-speed BCI system containing over 100 targets [20]. In 2021, Chen *et al* increased the number of BCI targets to 160. But the system's ITR was only 78.84 bits min^{-1} on average [21]. In 2022, Sun *et al* developed a 120-target BCI system with an average ITR of 265.74 bits min^{-1} [22].

It is worth noting that a counterbalance exists between the number of targets and the other two factors contributing to ITR improvement. In other words, as there are more targets, the BCI speed and accuracy are more likely to decrease, negatively impacting ITR. As a result, an elaborate coding strategy design is required to achieve a high ITR. On the other hand, the size of the monitor used to present the visual stimuli also affects the number of targets. Specifically, in previous studies, the target usually subtended 2°–4° of visual angle, and the space between two adjacent targets was 1°–5°. As a result, a 27-inch liquid-crystal display (LCD) screen with a 50° viewing angle could only display about 100 targets. Researchers may use a larger screen size to display more targets. For example, a 49-inch LCD was reported in a recent study [22]. However, a large screen is not practical for most conditions, and a bulky monitor hinders the practicality of real-life BCI applications. Alternatively, we could make the target smaller and closer together to fit more targets on a standard screen. However, a smaller target would result in weaker electroencephalography (EEG) features, and closer proximity of targets would cause more confusion in SSVEP classification. One remedy is to introduce hybrid EEG features to provide more useful information. But to the best of our knowledge, no study has yet implemented a high-speed BCI system with over 200 targets.

This study developed a novel paradigm that could effectively encode 216 targets using the hybrid features of P300, mVEP, and SSVEP. Specifically, all targets were divided into 36 sub-spellers encoded by different frequencies and each sub-speller contained six characters encoded in various time slots. Notably, this BCI system was deployed on a typical platform of a 27-inch LCD with 60 Hz and a 1920 × 1080 pixel resolution, which can be easily controlled by a laptop. In the decoding stage, the target identification required two sequential steps: sub-speller recognition and character recognition within the sub-speller. We can only identify the target character when both steps are performed correctly. The proposed BCI system achieved an average online ITR of 302.83 bits min^{-1}, with a peak value of 367.83 bits min^{-1}.

2. Materials and methods

2.1. Subjects
Ten healthy volunteers (four females and six males, 22–29 years of age, all right-handed) with normal or corrected to normal vision participated in both offline and online cued-guided spelling experiments. Another group of ten subjects participated in an online copy-spelling experiment. Four of them (i.e. S1, S3, S4, S7) also took part in the offline and online cued-guided spelling experiments (corresponding to S1, S7, S2, S3, respectively). The Institutional Review Board at Tianjin University approved the experimental protocol. According to the Helsinki Declaration, all subjects were fully informed of all procedures, understood all possible consequences of the study, and signed an informed consent form.

2.2. A hybrid mVEP-P300-SSVEP BCI paradigm
The visual stimuli were presented on a 27-inch LCD monitor with a 1920 × 1080 pixel resolution and a 60 Hz refresh rate. Figure 1(a) depicts a 12 × 18 matrix on a black background. The matrix was divided into an upper and lower half, comprising 36 6-character rectangles, also known as sub-spellers. The name of the 36 sub-spellers' first letter was given as an alphanumeric character (i.e. A–Z, 0–9). The six characters that make up each sub-speller are numbered 1, 2, …, 6 from top to bottom. Therefore, A1, A2, .., Z5, Z6, 01, 02, …, 95, 96 were used to identify the 216 stimulation targets. Each stimulation target has a vertical visual angle of 1.78 degrees and a horizontal angle of 2.32 degrees.

As shown in figure 1(b), the frequency division multiple access (FDMA) method was employed between sub-spellers, ranging from 10.4 Hz to 17.4 Hz with a 0.2 Hz interval. Empirically, the stimulation phase for FDMA was set to cycle from 0 to 2π with a step of 0.35π [12, 13]. As shown in figure 2, the time division multiple access (TDMA) method was adopted within sub-spellers. A random and ergodic colour change from black to yellow highlighted each of the six characters in the sub-speller individually. Meanwhile, a red vertical line with a height of 1.49° visual angle started moving leftward at a speed of 8.47° s^{-1} from the right edge of the characters to the left edge of the rectangle. The red vertical line produced a fleeting motion stimulus to elicit an mVEP response. The stimulus duration for each character was 200 ms, and the inter-stimulus interval was −100 ms. All sub-spellers were activated simultaneously and the flashing sequence between sub-spellers
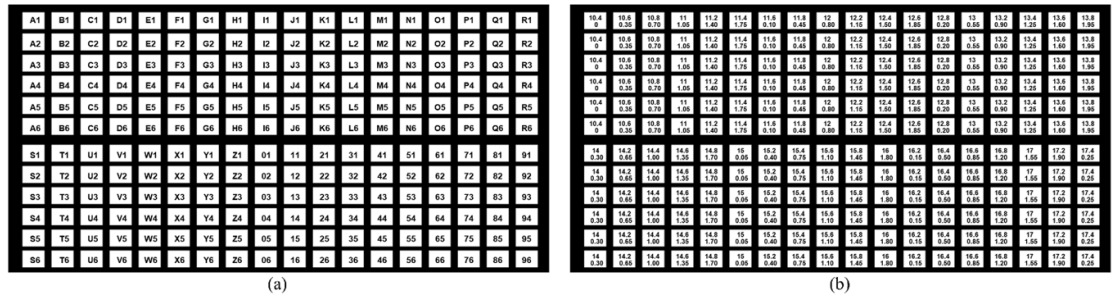
Figure 1. The display of the hybrid BCI speller. (a) The layout of the 216 commands. The interface was a 12×18 matrix divided into an upper and a lower half. (b) The stimulation frequency (Hz) and initial phase (π) for the 216 commands.
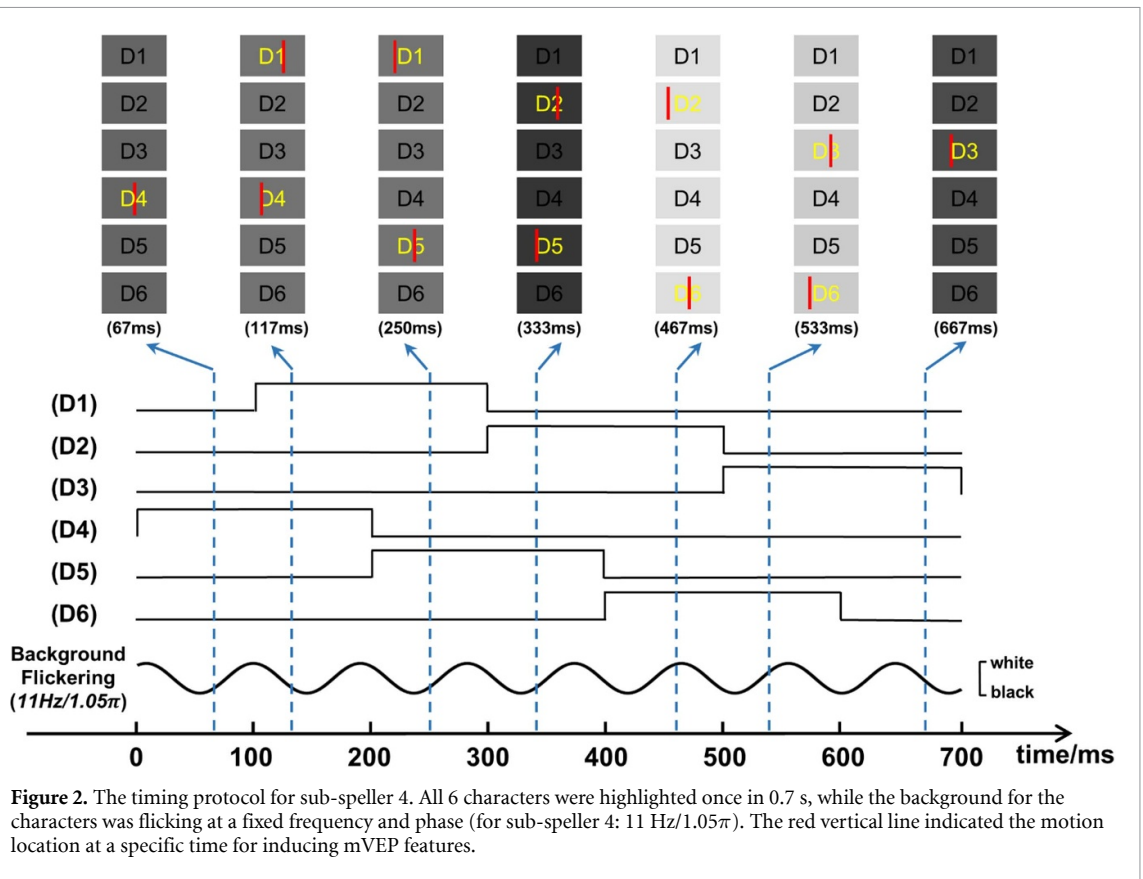

Figure 2. The timing protocol for sub-speller 4. All 6 characters were highlighted once in 0.7 s, while the background for the characters was flicking at a fixed frequency and phase (for sub-speller 4: 11 Hz/1.05π). The red vertical line indicated the motion location at a specific time for inducing mVEP features.

was the same. Thus, it only took 700 ms for all characters to flash once, called a round.

It should be noted that each stimulus rectangle flickered at a fixed frequency during the flashing of characters. In other words, there was a background flicking for each character gazed at by the subjects. Therefore, unlike the traditional hybrid paradigm design, the SSVEP feature lasted 0.7 s for a round, and the P300 feature was evoked once within a round. As previous studies reported [20, 23], the event-related potentials (ERPs) are mainly distributed at low frequencies, so the stimulation frequencies used in this study were purposefully set to be higher than 10 Hz. Based on this, a specific filter can be designed to separate the evoked mVEP-P300 and SSVEP, making feature extraction and identification easier.

2.3. BCI experiment

Participants sat on a chair in front of the screen at a distance of 60 cm. There were two experiments, namely offline experiments that were mainly used for feature analysis, algorithm comparison, and parameter tuning, and online experiments that were primarily used for validating the system. In the offline experiments, a red triangle with a 0.69° visual angle beneath the target character served as a cue for the target character's gaze for 0.5 s. Participants were instructed to concentrate on the target character during this time and count covertly how many times the target was highlighted (i.e. six times). The six rounds of flicking described above constitute a block. The participants needed to spell all 216 characters, which were divided into six groups presented

in random order. That is to say, each group contained 36 target characters that were selected by the stratified sampling method. The participants were allowed to take breaks between groups. It took about 17 min to complete the offline experiments.

In the online experiment, there were two types of online tasks (i.e. cued-guided spelling and free-spelling experiments). Classification algorithm and the other aspects of the two online experiments were completely consistent except for with or without visual cues. For online cued-guided spelling experiments, the participants were asked to spell all 216 characters in random sequences, which were also divided into six groups. The characters in each group were drawn together hierarchically. As same as the offline stage, the target character was indicated by an underneath red triangle for 0.5 s. For free-spelling experiment, another group of ten subjects were recruited. Before the experiment, the participants were trained to remember the position of each character and the character spelling sequences, e.g. A1–B2–C3–D4–E5–F6–G1–H2–…–73–85–96. Notably, each sequence includes 36 target characters and is arranged with a certain regularity in order to be able to calculate the accuracy of free spelling. A total of five character sequences (i.e. 180 target characters) were conducted. Participants were able to shift their fixation points very fast from character to character in the specific second without visual cues. And they could take a break between sequences. The duration time of visual shift for each participant was a fixed value, which was determined on the individual condition. In both online experiments, the flickering stimulation was performed in only one round.

2.4. EEG recording and processing

The Neuroscan Synamps2 system collected the EEG data from 64 electrodes in accordance with the international 10–20 system. The reference electrode was placed on the left mastoid, and the ground electrode was placed on the prefrontal lobe. The signals were sampled at 1000 Hz and saved on the computer after 0.1–200 Hz bandpass filtering and 50 Hz notch filtering.

In the target identification stage, there were two sequential steps: (a) recognizing the sub-speller containing the target character and then (b) recognizing the target character within the identified subspeller. The correct results can be obtained only when each step is correct. In offline analysis, six-fold cross-validation was used to evaluate the robustness of the classification accuracy. The sample allocation was conducted in the stratified sampling principle. It meant that 36 samples, one for each sub-speller, were used as the test set, and the remaining 180 (216−36 = 180) samples, five for each sub-speller, were used as the train set. In online analysis, all the data collected in the offline experiments, i.e. 216 samples, were used as the train set to train the

classification model. The corresponding specific data processing flow of mVEP-P300, and SSVEP features are introduced as follows, respectively.

For mVEP and P300 features, 16 EEG channels (F3/4, Fz, T7/8, C3/4, Cz, P3/4/7/8, Pz, PO7/8, Oz) were down-sampled to 250 Hz and filtered by the 0.5–10 Hz bandpass with Chebyshev type I filter. Then the sampling rate was down to 25 Hz and the mVEP-P300 feature was extracted from 50–800 ms.

For SSVEP feature, there were two channel montages for classification. The conventional montage of the nine occipital channels (Pz, POz, PO3/4, PO5/6, Oz, and O1/2) and a new montage of 30 parietal-occipital channels (CPz, CP1/2, CP3/4, CP5/6, TP7/8, Pz, P1/2, P3/4, P5/6, P7/8, POz, PO3/4, PO5/6, PO7/8, Oz, O1/2, and CB1/2) were first down-sampled to 250 Hz and filtered by a filter bank (including six Chebyshev type I filters) into [X Hz, 72 Hz] (X = 8, 18, 28, 38, 48, and 58). SSVEP features were extracted from 140–840 ms.

2.5. Ensemble task-related component analysis (TRCA)

The ensemble task-related component analysis (TRCA) has been proved the powerful recognition algorithm for SSVEP classification [18, 24]. TRCA is an algorithm that finds a projection matrix $w = \left[w_{j_1}, w_{j_2}, w_{j_3}, \ldots, w_{N_{ch}} \right]$ to maximize the covariance of task-related components between trials of the same class. j_1 and j_2 refer to the index of channels. N_{ch} is the number of channels. The specific process is as follows. First, assuming the recorded EEG signal $x(t) \in R^{N_{ch} \times N_P}$. N_p is the number of time points. Then, all possible combinations of inter-trials in projection space are summed as:

$$
w^T S w = \sum_{\substack{h_1, h_2 = 1 \\ h_1 \neq h_2}}^{N_t} \sum_{j_1, j_2 = 1}^{N_{ch}} w_{j_1} w_{j_2} \text{Cov}\left(x_{j_1}^{(h_1)}(t), x_{j_2}^{(h_2)}(t) \right)
$$

$$(1)$$

where, Cov(*a, b*) refers to the covariance between *a* and *b*. N_t is the number of training trials. h_1 and h_2 is the index of trial. The periods of $x_{j_1}^{(h)}(t)$ are fixed as t $\in$ [t_h, t_h + T]. Here, t_h is the beginning of the *h*-th trial and T is the duration of the *h*-th trial. To obtain the final results, the following restriction is termed:

$$
w^T Q w = \sum_{j_1, j_2 = 1}^{N_{ch}} w_{j_1} w_{j_2} \text{Cov}\left(x_{j_1}(t), x_{j_2}(t) \right) = 1. \quad (2)
$$

The finding of projection direction is transformed into an optimization problem:

$$
\hat{w} = \arg \max_w \left(w^T S w \right) / \left(w^T Q w \right). \quad (3)
$$

The equation can be solved by Lagrange multiplier method. Thus, the eigenvector of the matrix $Q^{-1}S$ is the optimal spatial direction. Finally, for N_f

stimulus frequencies, the ensemble TRCA (eTRCA) that integrated all the stimulus frequencies is constructed as:

$$W = \left[w_1, w_2, w_3, \ldots, w_{N_f} \right] \qquad (4)$$

2.6. Task-discriminant component analysis

Task-discriminant component analysis (TDCA) was proposed last year and achieved state-of-the-art performance [25]. In contrast to the strategy of TRCA that finds spatial space for each class, discriminant analysis finds the spatial-temporal filter of all classes. The details of TDCA are as follows. First, for each training trial, the EEG data were extended along time points:

$$\tilde{X} = \left[X^T, X_1^T, \ldots, X_l^T \right] \qquad (5)$$

where $X \in R^{N_{ch} \times N_p}$ refers to EEG data, equivalent to the $x(t)$ in TRCA. $\tilde{X} \in R^{(l+1)N_{ch} \times N_p}$ denotes the augmented EEG trial. $X_l \in R^{N_{ch} \times N_p}$ is obtained by delaying l points for X. If the delayed points exceed the original data length, a zero-padding operation is performed on the extended data. Notably, the same manner is carried out in the test trial. Second, the augmented EEG trial is projected onto the orthogonal space:

$$\tilde{X}_p = \tilde{X} P_i \qquad (6)$$

P is the orthogonal projection matrix of i th class after QR decomposition of the sine-cosine reference signal. The equations are:

$$Y_i = QR \qquad (7)$$

$$P_i = QQ^T \qquad (8)$$

The construction method of the reference signal Y_i refers to [13, 26]. Third, the secondary augmented EEG trial $X_a \in R^{(l+1)N_{ch} \times 2N_p}$ is formed as follows:

$$X_a = \left[\tilde{X}, \tilde{X} p \right] \qquad (9)$$

Finally, a two-dimensional linear discriminant analysis (LDA) is conducted to find the projection directions to discriminate trials from all classes. The between-class difference matrix $H_b \in R^{(l+1)N_{ch} \times 2N_cN_p}$ and the within-class difference matrix $H_w \in R^{(l+1)N_{ch} \times 2N_tN_p}$ are formed as:

$$H_b = \frac{1}{\sqrt{N_c}} \left[\bar{X}_a^{\,1} - \bar{X}_a^{\,a}, \ldots, \bar{X}_a^{\,N_c} - \bar{X}_a^{\,a} \right] \qquad (10)$$

$$H_w = \frac{1}{\sqrt{N_t}} \left[X_a^{(1)} - \overline{X}_a^{(1)}, \ldots, X_a^{(N_t)} - \overline{X}_a^{(N_t)} \right] \qquad (11)$$

where $\overline{X^i}$ and $\overline{X}^{(i)}$ refer to the two-dimensional class centres of i th class and the i th sample, respectively. N_c refers to the number of samples in Class c. The

superscript represents all classes, and $\overline{X_a}^a$ is calculated as follows:

$$\bar{X}_a^{\,a} = \frac{1}{N_t} \sum_{i=1}^{N_t} X_a^{(i)} \qquad (12)$$

According to the two-dimensional LDA method [27], the following Fisher criterion is constructed:

$$\hat{W} = \arg\max \frac{tr\left(W^T H_b H_b^T W \right)}{tr\left(W^T H_w H_w^T W \right)} \qquad (13)$$

where W denotes the projection direction. It also can be converted to an optimization problem and solved by Lagrange multiplier method as TRCA.

2.7. Various version of LDA

LDA is the most popular algorithm for P300-based BCIs because of its low computing resource requirements and simple implementation [28, 29]. LDA is a supervised dimensionality reduction method designed to increase between-class variance while reducing within-class variance. For a P300-based BCI, a set of EEG samples were referred to $x_i \in R^D (D = N_{ch}N_p) (i = 1, 2, \ldots, N_{all})$ and the corresponding class labels $C_i \in \{1, 2\}$. First, the means and empirical covariance of two classes are formed as follows:

$$\mu_C = \frac{1}{N_C} \sum_{i=1}^{N_c} x_i \qquad (14)$$

$$\Sigma_C = \frac{1}{N_C - 1} \sum_{i=1}^{N_c} (x_i - \mu_C)(x_i - \mu_C)^T, \ C = 1, 2 \qquad (15)$$

Based on the LDA theory, the equation can be formulated:

$$J(w) = \frac{w^T S_B w}{w^T S_W w} \qquad (16)$$

Where S_B and S_W are given by:

$$S_B = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T \qquad (17)$$

$$S_W = \frac{N_1}{N_{all}} \Sigma_1 + \frac{N_2}{N_{all}} \Sigma_2 . \qquad (18)$$

Thus, the overall goal is to maximize the $J(w)$ term. To find the optimal w, the term $J(w)$ is differentiated with respect to w, and set the derivatives equal to 0 to find the extremum. Finally, the solution is transferred to a generalized eigenvalue problem as follows:

$$S_B w = \lambda S_w w . \qquad (19)$$

The optimal projection is $w = S_W^{-1}(\mu_1 - \mu_2)$. LDA works well for binary problems like P300-based BCIs.

However, the LDA estimator may experience overfitting issues if the sample size is insufficient.

Some researchers developed extended algorithms based on the LDA theory, such as stepwise LDA (SWLDA) and shrinkage LDA (SKLDA). SWLDA selects the most significant features by combining forward and backward stepwise analyses, which can reduce the feature space [30]. The three parameters to be set are the maximum predetermined number of features and the *p*-values of adding and removing features. In this study, the *p*-value for entry was set to <0.1, while for removal was set to >0.15. The predetermined number of features was set to 60. Because there are high-dimensional features with only a few data samples, the estimated empirical covariance is imprecise, which hinders the estimator performance. SKLDA can solve this problem by modifying the extreme eigenvalues toward the average eigenvalue. Please see more details in [31]. The algorithms vectorise and concatenate the time points and multi-channels without spatial information. Spatial-temporal discriminant analysis (STDA) builds spatial-temporal two-way samples as feature matrices and performs alternating spatial and temporal optimization, which works well for insufficient samples [32].

2.8. Performance evaluation

To evaluate the performance of high-speed BCI, classification accuracy and ITR were used as evaluation indices in this study, which have been widely used in BCI research [1]. The ITR can be calculated as follows:

$$
\text{ITR} = \left[\log_2 N + P \log_2 P + (1-P) \log_2 \left(\frac{(1-P)}{N-1} \right) \right] \times \frac{60}{T}.
\tag{20}
$$

where N is the number of instruction sets, P is the classification accuracy, and T is the total amount of time needed to process each selection, i.e. cue time plus flashing time. In this study, the consuming times for 1–6 rounds were 1.2 s, 1.9 s, 2.6 s, 3.3 s, 4 s, and 4.7 s, respectively.

3. Results

3.1. EEG features analyses

In this study, mVEP, P300, and SSVEP features are induced to encode BCI targets. Figure 3 shows the grand average temporo-spatial patterns across all subjects, where the top row presents the ERP difference waveforms for all channels and their average, which were calculated by subtracting the non-target from the target waveforms, and the remaining rows show how the spatial pattern changed over time after target stimuli were presented. The amplitude in the occipital region was seen to gradually increase during the first 100 ms before decreasing. This might have relevance to the C1 component, which is primarily present near the calcarine fissure. During the 120–180 ms time
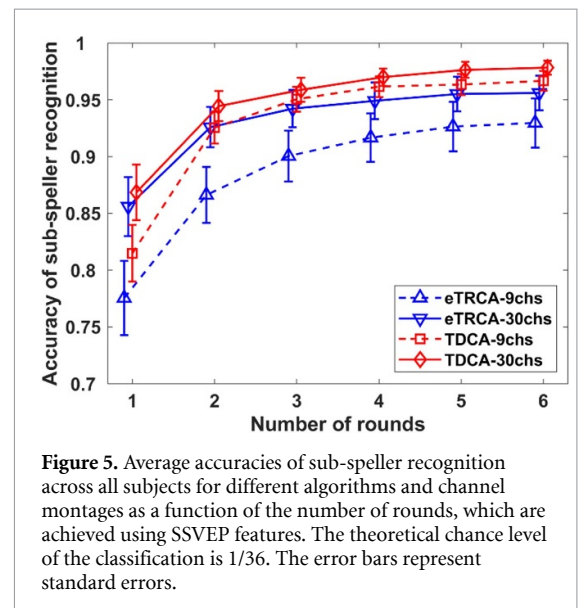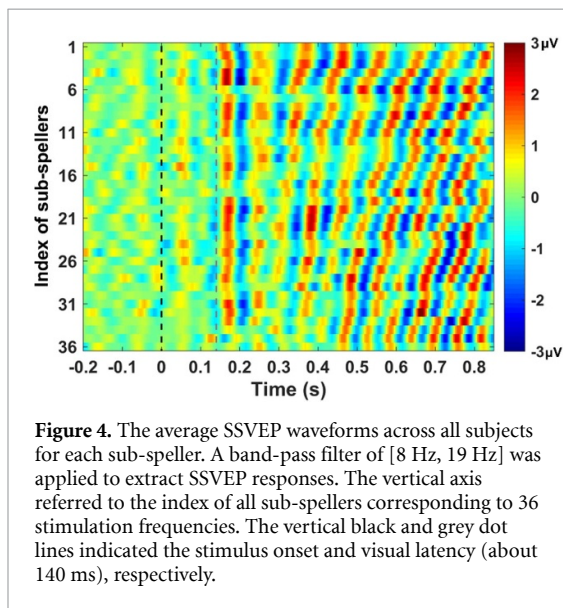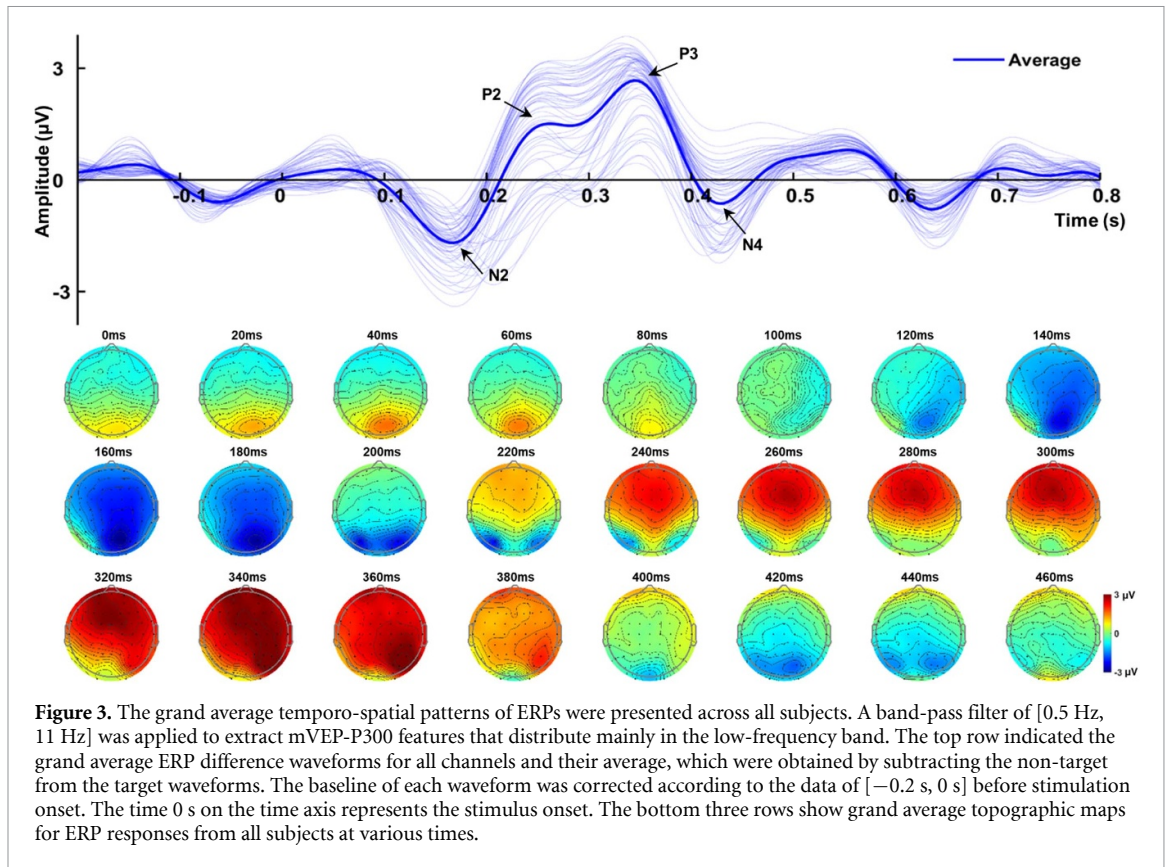
window, a negative deflection component was seen in the right occipital region due to the motion stimuli being located in the left visual field [6]. Its shifting distributions reflect the characteristics of mVEP. Then, the left occipital region of the EEG responses began to exhibit a negative deflection. A strong positive deflection of the P2 component was seen in the 240–300 ms time window with a parietal (up to central and frontal) topography. The distinct P300 component then started to emerge until about 380 ms. Finally, the bilateral occipital region exhibited the N4 component.

Figure 4 shows the average SSVEP waveforms at the fundamental frequency for each sub-speller. The distinct SSVEP responses for each sub-speller were observed after about 140 ms of stimulus onset, and there were differences among all sub-spellers. For example, the sub-speller #6 with the stimulation frequency of 11.4 Hz had approximately eight obvious periods (i.e. $11.4 \times 0.7 = 7.98$) while the sub-speller #31 with the stimulation frequency of 16.4 Hz had approximately 11 obvious periods (i.e. $16.4 \times 0.7 = 11.48$). The result demonstrated that the mVEP, P300, and SSVEP features were clearly evoked for this paradigm, which can be used for further classification.
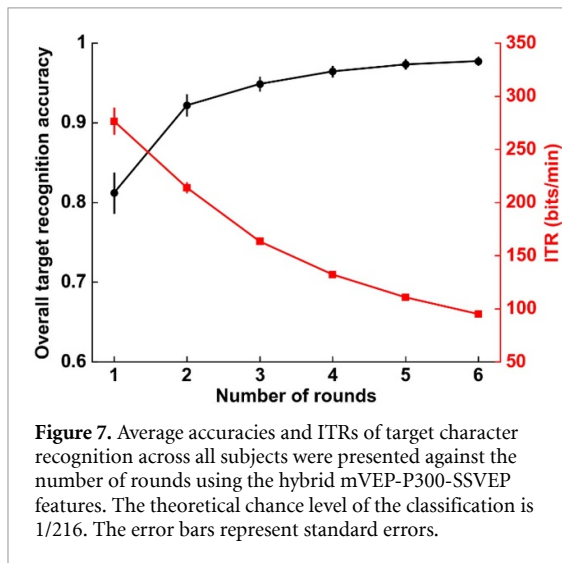
3.2. Offline BCI performance

As previously described, there were two steps in target identification: sub-speller recognition and character recognition within the sub-speller. Identification of the target character can only be achieved by correctly performing both steps. Figure 5 shows the average accuracies of sub-speller recognition against the number of rounds. The most powerful eTRCA and TDCA algorithms for detecting SSVEP features were used to evaluate the performance. Moreover, this figure also compares the performance of using two different EEG montages, namely the conventional montage of the nine occipital channels (i.e. Pz, POz, PO3/4, PO5/6, Oz, and O1/2) and a new montage of 30 parietal-occipital channels (i.e. CPz, CP1/2, CP3/4, CP5/6, TP7/8, Pz, P1/2, P3/4, P5/6, P7/8, POz, PO3/4, PO5/6, PO7/8, Oz, O1/2, and CB1/2). As a result, the STDA with 30 channels achieved the highest average accuracy at 1–6 rounds (86.85%, 94.44%, 95.88%, 96.99%, 97.64%, and 97.82%). As shown in the figure, the accuracy gradually improved as the number of rounds increased. Three-way repeated measures ANOVA revealed significant differences in accuracy for algorithms ($F(1, 9) = 8.15$, $p < 0.05$), montage ($F(1, 9) = 20.79$, $p < 0.01$), and the number of rounds ($F(5, 45) = 44.71$, $p < 0.001$). There were also significant interactions between algorithms and channel montage ($F(1, 9) = 6.92$, $p < 0.05$).

Figure 6 shows average accuracies of character recognition within sub-spellers across all subjects. The classic algorithms of LDA and its extended methods were adopted to compare system performance. As

Figure 3. The grand average temporo-spatial patterns of ERPs were presented across all subjects. A band-pass filter of [0.5 Hz, 11 Hz] was applied to extract mVEP-P300 features that distribute mainly in the low-frequency band. The top row indicated the grand average ERP difference waveforms for all channels and their average, which were obtained by subtracting the non-target from the target waveforms. The baseline of each waveform was corrected according to the data of [−0.2 s, 0 s] before stimulation onset. The time 0 s on the time axis represents the stimulus onset. The bottom three rows show grand average topographic maps for ERP responses from all subjects at various times.

Figure 4. The average SSVEP waveforms across all subjects for each sub-speller. A band-pass filter of [8 Hz, 19 Hz] was applied to extract SSVEP responses. The vertical axis referred to the index of all sub-spellers corresponding to 36 stimulation frequencies. The vertical black and grey dot lines indicated the stimulus onset and visual latency (about 140 ms), respectively.

Figure 5. Average accuracies of sub-speller recognition across all subjects for different algorithms and channel montages as a function of the number of rounds, which are achieved using SSVEP features. The theoretical chance level of the classification is 1/36. The error bars represent standard errors.

expected from the SSVEP results, accuracy tended to increase with the number of rounds. All algorithms achieved an accuracy above 90% regardless of the number of rounds. For example, at one round, the algorithms' accuracy for LDA, SWLDA, SKLDA, and STDA was 92.31%, 91.06%, 91.81%, and 90.97%, respectively. Two-way repeated measures ANOVA showed significant main effects of algorithms ($F_{(3, 27)} = 4.66$, $p < 0.01$) and the number of rounds ($F_{(5, 45)} = 29.86$, $p < 0.001$) on accuracy. The algorithms and the number of rounds had significant interactions ($F_{(15, 135)} = 2.18$, $p < 0.01$). By post-hoc pairwise comparison using paired t-test, the LDA outperformed SWLDA ($t_9 = 2.49$, $p < 0.05$) and STDA ($t_9 = 2.82$, $p < 0.05$). The ratio of the sample to the feature vector dimension in the train set was 1080:320 for LDA, SKLDA, 1080:60 for SWLDA, 1080:16, or 1080:200 for STDA, suggesting that there was no issue with insufficient sample size for this study. Except for LDA, these algorithms were proposed for solving insufficient data problems to mitigate overfitting [30–32]. In conclusion, the

Figure 6. Average accuracies of character recognition within sub-spellers across all subjects for different algorithms are plotted against the number of rounds utilizing mVEP-P300 features. The theoretical chance level of the classification is 1/6. The error bars show standard errors.

Figure 7. Average accuracies and ITRs of target character recognition across all subjects were presented against the number of rounds using the hybrid mVEP-P300-SSVEP features. The theoretical chance level of the classification is 1/216. The error bars represent standard errors.

TDCA with 30 channels for SSVEP features and the LDA for mVEP-P300 features achieved the highest accuracy, so they were chosen to recognize the target character next.

Figure 7 shows the average accuracies and ITRs of target character recognition. The hybrid mVEP-P300-SSVEP features achieved accuracy of 81.19%, 92.18%, 94.86%, 96.44%, 97.31%, and 97.73% at 1–6 rounds, respectively. The corresponding average ITR achieved the maximum value of 276.52 bits min^{-1} at 1 round. Therefore, only one round was used for the online test.

3.3. Online BCI performance

Table 1 lists the results of the online cued-guided spelling BCI experiments across subjects. A fixed stimulus duration of one round was used in the online experiment. The average accuracy across subjects was 85.37% ± 7.49%, resulting in an average ITR of 302.83 ± 39.20 bits min^{-1}. Notably, the maximal and minimal ITRs were 367.83 bits min^{-1} (S4) and 248.17 bits min^{-1} (S3), respectively.

Table 2 lists the results of the online free-spelling experiment. The duration time of visual shift for each subject was a fixed value, which was determined by the individual condition. The mean accuracy was 86.00% and the mean ITR was 204.47 bits min^{-1} across all subjects. Due to the increase of visual shift time, the mean spelling rate of 32.97 characters per minute was much lower than 50 characters per minute in the cued-guided spelling experiment. These results demonstrated the feasibility and effectiveness of the proposed high-speed hybrid BCI system with over 200 targets.

4. Discussion

4.1. Advantages of the hybrid P300-mVEP-SSVEP features

From the perspective of communication, the encoding methods of BCIs could be divided into TDMA, FDMA, code division multiple access, and space division multiple access [3]. Among them, TDMA and FDMA are one of the most widely studied methods, and the typical representative features are P300 and SSVEP features. TDMA divides multiple targets into different time slots. Due to the characteristics of TDMA, although an unlimited number of targets can be achieved theoretically, the system performance would be severely hurt and cannot meet the requirements of practical applications. FDMA divides multiple targets into different frequency bands. But the available frequency bands are relatively narrow, resulting in a limited number of targets. The inherent limitations of a single encoding method make it difficult to implement a high-speed BCI system with a large instruction set. Therefore, a combination of several BCI paradigms may overcome their shortcomings and expand the instruction set in an efficient manner.

The hybrid BCIs combine two or more different brain activity patterns or input signal sources, which have been demonstrated to have irreplaceable advantages [11, 20]. For example, hybrid BCIs can combine different EEG features (e.g. P300-SSVEP) [30], EEG and electromyogram [33], or EEG and electro-oculogram [34], etc. There are two ways to combine things, serial and parallel modes. The serial mode is similar to the 'and' operation in logical operations, where the correct results can only be obtained if each step is correct. The parallel mode resembles the 'or' operation in logical operations, where the identification of each feature impacts the final results. Traditional hybrid BCIs are more often developed using two features and have made significant strides [12, 20, 30]. Panicker *et al* implemented a 36-target BCI speller by combining SSVEP as a brain switch with P300 features in the serial mode [11]. However, to achieve a higher performance BCI system,

Table 1. Results of online cued-guided spelling experiments.

Subject	Consuming time (s)	Selections (Correct/Total)	Accuracy (%)	ITR (bits min^{-1})
S1	1.2 (0.5 + 0.7)	204/216	94.44	350.74
S2	1.2 (0.5 + 0.7)	181/216	83.80	293.01
S3	1.2 (0.5 + 0.7)	161/216	74.54	248.17
S4	1.2 (0.5 + 0.7)	210/216	97.22	367.83
S5	1.2 (0.5 + 0.7)	186/216	86.11	304.87
S6	1.2 (0.5 + 0.7)	170/216	78.70	267.89
S7	1.2 (0.5 + 0.7)	178/216	82.41	286.03
S8	1.2 (0.5 + 0.7)	163/216	75.46	252.49
S9	1.2 (0.5 + 0.7)	190/216	87.96	314.59
S10	1.2 (0.5 + 0.7)	201/216	93.06	342.65
Max	—	—	97.22	367.83
Min	—	—	74.54	248.17
Mean ± STD	—	—	**85.37 ± 7.49**	**302.83 ± 39.20**

Table 2. Results of online free-spelling experiments.

Subject	Consuming time (s)	Selections (Correct/Total)	Accuracy (%)	ITR (bits min^{-1})
S1	1.5 (0.8 + 0.7)	171/180	95.00	283.24
S2	1.7 (1.0 + 0.7)	166/180	92.22	238.52
S3	1.9 (1.2 + 0.7)	153/180	85.00	188.93
S4	1.8 (1.1 + 0.7)	155/180	86.11	203.25
S5	2.1 (1.4 + 0.7)	151/180	83.89	167.70
S6	1.7 (1.0 + 0.7)	162/180	90.00	229.80
S7	1.6 (0.9 + 0.7)	148/180	82.22	213.83
S8	2.2 (1.5 + 0.7)	163/180	90.56	179.24
S9	1.9 (1.2 + 0.7)	141/180	78.33	168.07
S10	1.8 (1.1 + 0.7)	138/180	76.67	172.11
Max	—	—	95.00	283.24
Min	—	—	76.67	167.70
Mean ± STD	—	—	**86.00 ± 5.98**	**204.47 ± 37.56**

more feature patterns may be needed to provide more useful information. This study combined three kinds of EEG features, i.e. P300, mVEP, and SSVEP, to develop a high-speed hybrid BCI with more than 200 targets. The TDMA and FDMA methods were combined in an elaborate manner. Specifically, the P300 and mVEP features were induced simultaneously and combined in parallel mode. During the classification stage, these two features were extracted simultaneously to recognize the target character within a specific sub-speller. Moreover, the above two features were combined in serial mode with SSVEP, which was used to recognize sub-spellers. The evoked SSVEP lasting 0.7 s offered enough high accuracies. As a result, using three features to encode BCI targets could provide more useful information and achieve a high-performance BCI system.

4.2. Compared with previous studies featuring large instruction sets

High-speed BCI systems with a large instruction set can broaden their application scenarios,

thereby attracting extensive research interest from researchers. Researchers have tackled this issue in various aspects, including encoding strategy [9, 22] and decoding method [13, 26]. For example, Xu *et al* proposed the concurrent P300 and SSVEP features containing four types of EEG features to develop a high-speed BCI system with over 100 targets [20]. Chen *et al* adopted the multiple frequency sequential coding method to achieve a BCI system with 160 targets [21]. Sun *et al* used the characteristics of code-modulated visual evoked potentials to reduce the training time of BCIs [22]. Here, we compared this work with the previous online BCI studies since the BCI was first proposed in 1973 [9, 11, 13, 18–22, 26, 30, 35–54]. It is worth noting that since not all studies adopted free-spelling experiment, we used the results of online cued-guided spelling experiment for comparison. As shown in figure 8, the number of targets and ITR for each study are indicated by a solid dot. Most studies had fewer than 100 targets and an ITR of lower than 100 bits min^{-1}. The *X* marks the average number of targets and
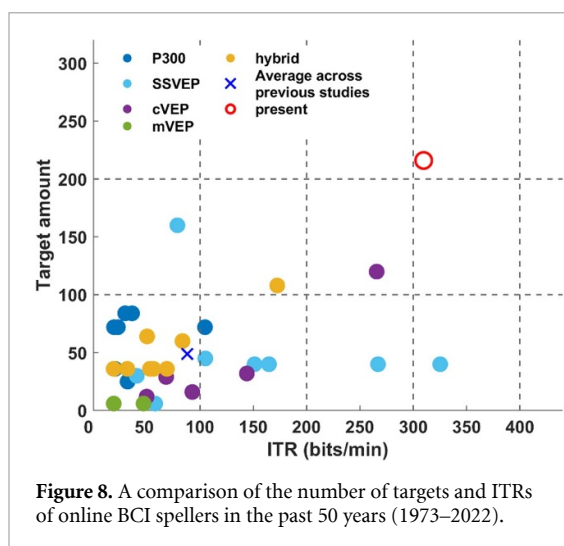
Figure 8. A comparison of the number of targets and ITRs of online BCI spellers in the past 50 years (1973–2022).

ITR across previous studies, which was 48.87 commands at 88.17 bits min^{-1}. Overall, the proposed hybrid BCI system had 3.4 times higher ITRs and 4.4 times more targets than the average over the previous 50 years.

5. Conclusion

This study used hybrid P300, mVEP, and SSVEP features to implement a high-speed BCI system with up to 216 targets. A novel hybrid paradigm was proposed to mark all characters with P300 and mVEP encoded in various time slots and SSVEP encoded by different frequencies. The TDCA and LDA algorithms were used to classify the hybrid features. As a result, the online cued-guided spelling experiment could reach an average accuracy of 85.37%, resulting in an average ITR of 302.83 bits min^{-1} and a maximum ITR of 367.83 bits min^{-1}. This study developed the first high-speed BCI system with more than 200 targets, which may lead to many new BCI application scenarios.

Data availability statement

The data generated and/or analysed during the current study are not publicly available for legal/ethical reasons but are available from the corresponding author on reasonable request.

Acknowledgments

ORCID iDs

Jin Han ⓘ https://orcid.org/0000-0002-6912-1738
Xiaolin Xiao ⓘ https://orcid.org/0000-0002-3516-561X

References

 [1] Wolpaw J R, Birbaumer N, Heetderks W J, McFarland D J, Peckham P H, Schalk G, Donchin E, Quatrano L A, Robinson C J and Vaughan T M 2000 Brain-computer interface technology: a review of the first international meeting *IEEE Trans. Rehabil. Eng.* **8** 164–73
 [2] Xu M, He F, Jung T P, Gu X and Ming D 2021 Current challenges for the practical application of electroencephalography-based brain–computer interfaces *Engineering* **7** 1710–2
 [3] Gao S, Wang Y, Gao X and Hong B 2014 Visual and auditory brain–computer interfaces *IEEE Trans. Biomed. Eng.* **61** 1436–47
 [4] Wang K, Xu M, Wang Y, Zhang S, Chen L and Ming D 2020 Enhance decoding of pre-movement EEG patterns for brain–computer interfaces *J. Neural Eng.* **17** 016033
 [5] Guo F, Hong B, Gao X and Gao S 2008 A brain–computer interface using motion-onset visual evoked potential *J. Neural Eng.* **5** 477
 [6] Chen J, Li Z, Hong B, Maye A, Engel A K and Zhang D 2018 A single-stimulus, multitarget BCI based on retinotopic mapping of motion-onset VEPs *IEEE Trans. Biomed. Eng.* **66** 464–70
 [7] Cattan G H, Andreev A, Mendoza C and Congedo M 2019 A comparison of mobile VR display running on an ordinary smartphone with standard PC display for P300-BCI stimulus presentation *IEEE Trans. Games* **13** 68–77
 [8] Fernández-Rodríguez Á, Velasco-Álvarez F, Medina-Juliá M T and Ron-Angevin R 2019 Evaluation of emotional and neutral pictures as flashing stimuli using a P300 brain–computer interface speller *J. Neural Eng.* **16** 056024
 [9] Chen Y, Yang C, Chen X, Wang Y and Gao X 2021 A novel training-free recognition method for SSVEP-based BCIs using dynamic window strategy *J. Neural Eng.* **18** 036007
[10] Ge S, Jiang Y, Zhang M, Wang R, Iramina K, Lin P, Leng Y, Wang H and Zheng W 2021 SSVEP-based brain-computer interface with a limited number of frequencies based on dual-frequency biased coding *IEEE Trans. Neural Syst. Rehabil. Eng.* **29** 760–9
[11] Panicker R C, Puthusserypady S and Sun Y 2011 An asynchronous P300 BCI with SSVEP-based control state detection *IEEE Trans. Biomed. Eng.* **58** 1781–8
[12] Han J, Liu C, Chu J, Xiao X, Chen L, Xu M and Ming D 2022 Effects of inter-stimulus intervals on concurrent P300 and SSVEP features for hybrid brain-computer interfaces *J. Neurosci. Methods* **372** 109535
[13] Chen X, Wang Y, Nakanishi M, Gao X, Jung T P and Gao S 2015 High-speed spelling with a noninvasive brain–computer interface *Proc. Natl Acad. Sci. USA* **112** E6058–67
[14] Edelman B J, Meng J, Suma D, Zurn C, Nagarajan E, Baxter B, Cline C C and He B 2019 Noninvasive neuroimaging enhances continuous neural tracking for robotic device control *Sci. Robot.* **4** eaaw6844
[15] Vidal J J 1973 Toward direct brain-computer communication *Annu. Rev. Biophys. Bioeng.* **2** 157–80
[16] Farwell L A and Donchin E 1988 Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials *Electroencephalogr. Clin. Neurophysiol.* **70** 510–23
[17] Calhoun G L and McMillan G R 1996 EEG-based control for human-computer interaction *Proc. 3rd Annual Symp. on Human Interaction with Complex Systems HICS'96* pp 4–9
[18] Nakanishi M, Wang Y, Chen X, Wang Y T, Gao X and Jung T P 2017 Enhancing detection of SSVEPs for a

high-speed brain speller using task-related component analysis *IEEE Trans. Biomed. Eng.* **65** 104–12

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

[20] Xu M, Han J, Wang Y, Jung T P and Ming D 2020 Implementing over 100 command codes for a high-speed hybrid brain-computer interface using concurrent P300 and SSVEP features *IEEE Trans. Biomed. Eng.* **67** 3073–82

[21] Chen Y, Yang C, Ye X, Chen X, Wang Y and Gao X 2021 Implementing a calibration-free SSVEP-based BCI system with 160 targets *J. Neural Eng.* **18** 046094

[22] Sun Q, Zheng L, Pei W, Gao X and Wang Y 2022 A 120-target brain-computer interface based on code-modulated visual evoked potentials *J. Neurosci. Methods* **375** 109597

[23] Jin J, Daly I, Zhang Y, Wang X and Cichocki A 2014 An optimized ERP brain–computer interface based on facial expression changes *J. Neural Eng.* **11** 036004

[24] Tanaka H, Katura T and Sato H 2013 Task-related component analysis for functional neuroimaging and application to near-infrared spectroscopy data *NeuroImage* **64** 308–27

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

[26] Chen X, Wang Y, Gao S, Jung T-P and Gao X 2015 Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain–computer interface *J. Neural Eng.* **12** 046008

[27] Liao X, Yao D, Wu D and Li C 2007 Combining spatial filters for the classification of single-trial EEG in a finger movement task *IEEE Trans. Biomed. Eng.* **54** 821–31

[28] Ju J, Feleke A G, Luo L and Fan X 2022 Recognition of drivers' hard and soft braking intentions based on hybrid brain-computer interfaces *Cyborg Bionic Syst.* **2022** 9847652

[29] Li S, Jin J, Daly I, Liu C and Cichocki A 2022 Feature selection method based on Menger curvature and LDA theory for a P300 brain–computer interface *J. Neural Eng.* **18** 066050

[30] Xu M, Chen L, Zhang L, Qi H, Ma L, Tang J, Wan B and Ming D 2014 A visual parallel-BCI speller based on the time–frequency coding strategy *J. Neural Eng.* **11** 026014

[31] Blankertz B, Lemm S, Treder M, Haufe S and Müller K-R 2011 Single-trial analysis and classification of ERP components—a tutorial *NeuroImage* **56** 814–25

[32] Zhang Y, Zhou G, Zhao Q, Jin J, Wang X and Cichocki A 2013 Spatial-temporal discriminant analysis for ERP-based brain-computer interface *IEEE Trans. Neural Syst. Rehabil. Eng.* **21** 233–43

[33] Yang Q, Siemionow V, Yao W, Sahgal V and Yue G H 2010 Single-trial EEG-EMG coherence analysis reveals muscle fatigue-related progressive alterations in corticomuscular coupling *IEEE Trans. Neural Syst. Rehabil. Eng.* **18** 97–106

[34] Lee M H, Williamson J, Won D O, Fazli S and Lee S W 2018 A high performance spelling system based on EEG-EOG signals with visual feedback *IEEE Trans. Neural Syst. Rehabil. Eng.* **26** 1443–59

[35] Ryan D B, Frye G, Townsend G, Berry D, Mesa-G S, Gates N A and Sellers E W 2010 Predictive spelling with a P300-based brain–computer interface: increasing the rate of communication *Int. J. Hum.-Comput. Interact.* **27** 69–84

[36] Townsend G, LaPallo B K, Boulay C B, Krusienski D J, Frye G, Hauser C, Schwartz N E, Vaughan T M, Wolpaw J R and Sellers E W 2010 A novel P300-based brain–computer interface stimulus presentation paradigm: moving beyond rows and columns *Clin. Neurophysiol.* **121** 1109–20

[37] Jin J, Allison B Z, Sellers E W, Brunner C, Horki P, Wang X and Neuper C 2011 An adaptive P300-based control system *J. Neural Eng.* **8** 036006

[38] Jin J, Sellers E W and Wang X 2012 Targeting an efficient target-to-target interval for P300 speller brain–computer interfaces *Med. Biol. Eng. Comput.* **50** 289–96

[39] Hwang H J, Lim J H, Jung Y J, Choi H, Lee S W and Im C H 2012 Development of an SSVEP-based BCI spelling system adopting a QWERTY-style LED keyboard *J. Neurosci. Methods* **208** 59–65

[40] Bin G, Gao X, Yan Z, Hong B and Gao S 2009 An online multi-channel SSVEP-based brain–computer interface using a canonical correlation analysis method *J. Neural Eng.* **6** 046002

[41] Chen X, Chen Z, Gao S and Gao X 2014 A high-itr ssvep-based bci speller *Brain Comput. Interfaces* **1** 181–91

[42] Thielen J, Marsman P, Farquhar J and Desain P 2021 From full calibration to zero training for a code-modulated visual evoked potentials for brain–computer interface *J. Neural Eng.* **18** 056007

[43] Riechmann H, Finke A and Ritter H 2015 Using a cVEP-based brain-computer interface to control a virtual agent *IEEE Trans. Neural Syst. Rehabil. Eng.* **24** 692–9

[44] Bin G, Gao X, Wang Y, Hong B and Gao S 2009 VEP-based brain-computer interfaces: time, frequency, and code modulations *IEEE Comput. Intell. Mag.* **4** 22–26

[45] Spüler M, Rosenstiel W and Bogdan M 2012 Online adaptation of a c-VEP brain-computer interface (BCI) based on error-related potentials and unsupervised learning *PLoS One* **7** e51077

[46] Käthner I, Kübler A and Halder S 2015 Rapid P300 brain-computer interface communication with a head-mounted display *Front. Neurosci.* **9** 00207

[47] Lin Z, Zhang C, Zeng Y, Tong L and Yan B 2018 A novel P300 BCI speller based on the Triple RSVP paradigm *Sci. Rep.* **8** 3350

[48] Ma T, Li H, Yang H, Lv X, Li P, Liu T, Yao D and Xu P 2017 The extraction of motion-onset VEP BCI features based on deep learning and compressed sensing *J. Neurosci. Methods* **275** 80–92

[49] Liu T, Goldberg L, Gao S and Hong B 2010 An online brain–computer interface using non-flashing visual evoked potentials *J. Neural Eng.* **7** 036003

[50] Yin E, Zhou Z, Jiang J, Chen F, Liu Y and Hu D 2013 A novel hybrid BCI speller based on the incorporation of SSVEP into the P300 paradigm *J. Neural Eng.* **10** 026012

[51] Yin E, Zhou Z, Jiang J, Chen F, Liu Y and Hu D 2013 A speedy hybrid BCI spelling approach combining P300 and SSVEP *IEEE Trans. Biomed. Eng.* **61** 473–83

[52] Chang M H, Lee J S, Heo J and Park K S 2016 Eliciting dual-frequency SSVEP using a hybrid SSVEP-P300 BCI *J. Neurosci. Methods* **258** 104–13

[53] Yin E, Zeyl T, Saab R, Chau T, Hu D and Zhou Z 2015 A hybrid brain–computer interface based on the fusion of P300 and SSVEP scores *IEEE Trans. Neural Syst. Rehabil. Eng.* **23** 693–701

[54] Lin K, Cinetto A, Wang Y, Chen X, Gao S and Gao X 2016 An online hybrid BCI system based on SSVEP and EMG *J. Neural Eng.* **13** 026020