Effects of inter-stimulus intervals on concurrent P300 and SSVEP features for hybrid brain-computer interfaces

Jin Han [a], Chuan Liu [b], Jiayue Chu [b], Xiaolin Xiao [a,c,*], Long Chen [c,*], Minpeng Xu [a,c], Dong Ming [a,c]

[a] Department of Biomedical Engineering, College of Precision Instruments and Optoelectronics Engineering, Tianjin University, Tianjin 300072, People's Republic of China
[b] Division of Medicine, Tianjin University, Tianjin 300072, People's Republic of China
[c] Academy of Medical Engineering and Translational Medicine, Tianjin University, Tianjin 300072, People's Republic of China

ARTICLE INFO

ABSTRACT

Background: Recently, we have implemented a high-speed brain-computer interface (BCI) system with a large instruction set using the concurrent P300 and steady-state visual evoked potential (SSVEP) features (also known as hybrid features). However, it remains unclear how to select inter-stimulus interval (ISI) for the proposed BCI system to balance the encoding efficiency and decoding performance.

New method: This study developed a 6 * 9 hybrid P300-SSVEP BCI system and investigated a series of ISIs ranged from −175–0 ms with a step of 25 ms. The influence of ISI on the hybrid features was analyzed from several aspects, including the amplitude of the induced features, classification accuracy, information transfer rate (ITR). Twelve naive subjects were recruited for the experiment.

Results: The results showed the ISI factor had a significant impact on the hybrid features. Specifically, as the values of ISI decreased, the amplitudes of the induced features and accuracies decreased gradually, while the ITRs increased rapidly. It's achieved the highest ITR of 158.50 bits/min when ISI equal to − 175 ms.

Comparison with Existing Method: The optimal ISI in this study achieved superior performance in comparison with the one we used in the previous study.

Conclusions: The ISI can exert an important influence on the P300-SSVEP BCI system and its optimal value is − 175 ms in this study, which is significant for developing the high-speed BCI system with larger instruction sets in the future.

1. Introduction

Brain-computer Interface (BCI) is a system by detecting central nervous system activities and transforming it into manual output, which can make the brain directly communicate with the external environment (Wolpaw et al., 2000; Wolpaw et al., 2002; Feng et al., 2018; Wang et al., 2020). The visual BCI (v-BCI) system is an application of visual pathway in response to external stimuli that has the advantages of stable evoked features and fast speed (Gao et al., 2014; Xu et al., 2021). The typical features, such as P300 (Townsend and Platsko, 2016; Xiao et al., 2019), steady-state visual evoked potential (SSVEP) (Wang et al., 2006; Wang et al., 2008), and their hybrid features, have become increasingly popular paradigms (Panicker et al., 2011; Wang et al., 2015; Xu et al., 2014).

The large instruction set with high-speed has always been one of the important development directions for BCI community. Researchers have made a lot of efforts from two aspects, namely encoding strategy and decoding algorithm (Xu et al., 2014; Xu et al., 2021). For encoding strategy, Jin et al. designed an adaptive P300 BCI with 84 characters (Jin et al., 2011). Townsend et al. realized a 72 characters BCI system by designing a novel P300 checkerboard paradigm (CBP) (Townsend et al., 2010). Xu et al. developed a novel 36 commands BCI speller with information transfer rate (ITR) of 63.33 bits/min by miniature

---

  * Corresponding author at: Department of Biomedical Engineering, College of Precision Instruments and Optoelectronics Engineering, Tianjin University, Tianjin 300072, People's Republic of China.
  ** Corresponding author.
    E-mail addresses: xiaoxiao0@tju.edu.cn (X. Xiao), cagor@tju.edu.cn (L. Chen).

event-related potentials (ERPs), which can significantly reduce visual fatigue (Xu et al., 2018). Chen et al. implemented a 160 commands BCI system with average ITR of 78.84 bits/min using multiple frequency sequential coding (MFSC) method (Chen et al., 2021). For decoding algorithm, Chen et al. proposed the filter-bank method and extended CCA method successively to decode SSVEP BCI system with 40 commands (Chen et al., 2015a; Chen et al., 2015b). Nakanishi et al. proposed task-related component analysis (TRCA) that is the most powerful algorithm for SSVEP-based BCI (Nakanishi et al., 2017). Chen et al. developed a SSVEP-based BCI system used the dynamic window strategy with average online ITR of 164.72 bits/min, whereas the number of instruction was only 40 commands (Chen et al., 2021). The large instruction set BCI with high-speed had not been developed, mainly due to mutual restriction between the number of instructions and the consuming time of outputting a command. Its realization required highly sophisticated encoding design and advanced decoding algorithm. Recently, we have proposed the concurrent P300 and SSVEP features (also known as hybrid features in the context) and implemented an over 100 commands BCI system covering full keyboard keys with a maximum ITR of 238.41 bits/min (Xu et al., 2020). Nevertheless, the proposed BCI system has room in encoding strategy for further improvement to better meet the actual application.

The encoding performance of BCI was determined by encoding efficiency, which is mainly affected by three parameters: stimulus duration (SD), inter-stimulus interval (ISI), and stimulus onset asynchrony (SOA), respectively (Allison and Pineda, 2006; Polprasert et al., 2013). Actually, SOA is equal to SD plus ISI according to their definition. In order to quantify the relative relationship between ISI and SD, we defined the ratio of the opposite value of ISI to SD that was called as ROIS. It's worth noting that there is an overlap between the data corresponding to adjacent trials when ISI is negative value (i.e. ROIS is positive value). And the larger opposite value of ISI (or larger ROIS value), the more data overlap. From a view of pattern recognition, if the degree of overlap is increased, the encoding time would be shorter and the system response speed would be faster, but it would be brought greater difficulties for feature extraction and recognition. The exact impact on BCI system performance of ITR is not clear yet. If the degree of overlap is decreased, feature recognition will become relative easier, but the encoding time will be increased so that system performance of ITR will be also changed. On the other hand, based on the study of ERP (Luck, 2014; Martens et al., 2009), the changes of ISI or ROIS may have an impact on the amplitudes of the induced features. For the proposed hybrid features, besides the frequency-phase-modulated SSVEP and time-modulated P300 features as contained in the traditional hybrid P300 and SSVEP features, there are two new distinct EEG features, i.e. time-modulated SSVEP and frequency-phase-modulated P300 (Xu et al., 2020; Wang et al., 2016). More importantly, the effects of the encoding parameters on the hybrid features and system performance are not clear yet, which is one of the key issues to be explored urgently to enhance the performance of high-speed BCI with a large instruction set.

This study aimed to explore the effects of ISI on the concurrent P300 and SSVEP features and the BCI system performance. We designed a 6 * 9 hybrid P300-SSVEP BCI system, whose SD was a fixed value of 200 ms and ISI was a variable value ranged from 0 ms to − 175 ms with a step of − 25 ms. Then the influence of ISI on P300 and SSVEP features was studied separately from the several aspects of the amplitude of the induced features, classification accuracy, ITR, and so on. The results showed that the hybrid features achieved the best performance compared to traditional SSVEP or SSVEP+P300 features. And the ISI have a significant impact on hybrid features. Finally, it's achieved the highest ITR of 158.50 bits/min when ISI equal to − 175 ms, which is the optimal and recommended setting. For the same BCI paradigm design, the ITRs in this study have an impressive improvement using optimal parameter setting, which provides an effective way for developing high-speed BCI with larger instruction sets and broaden application prospects.

2. Materials and methods

2.1. Subjects

Twelve healthy volunteers (five females and seven males) aged between 20 and 21 with normal or corrected vision took part in the experiment. According to the Helsinki Declaration, all subjects were fully aware of all procedures, understood all possible consequences of the study, and signed an informed consent form. The study was approved by the Review Committee of Tianjin University. Before participating in the experiment, all the subjects had not experienced the relevant type of experimental training.

### 2.2. Hybrid P300-SSVEP BCI speller paradigm

The stimulation interface was presented on a 27-inch liquid crystal display (LCD) with a resolution of 1920 * 1080 pixels and a refresh rate of 120 Hz. A 6 * 9 matrix showed 54 black characters on a white background, as shown in the Fig. 1a. They were further divided into 6 independent small 3 * 3 character matrices, which were also named as P300 sub-speller. This is because that the 9 characters in the sub-speller were encoded by time division multiple access (TDMA) and were individually highlighted by a gray square in a pseudo-random and ergodic sequence. Each stimulation square subtended 3.33 degrees of visual angle in the vertical direction and 3.33 degrees in the horizontal direction. The stimulus duration for each character was 200 ms.

The 6 sub-spellers were activated simultaneously. A complete traversal flashing of 54 characters is called as a round. The hybrid encoding scheme was used in this study. For sub-spellers, frequency division multiple access (FDMA) was adopted. In order to distinguish the frequency band of SSVEP and P300, the flicking frequencies were greater than 10 Hz and ranged from 11 Hz to 16 Hz with 1 Hz interval. And the initial phase of the six sub-spellers was also different. According to (Wang et al., 2016), SSVEPs with any stimulation phase for 40 frequencies (ranging from 8 Hz to 15.8 Hz with step of 0.2 Hz), which covered all the stimulation frequencies of this experiment, can be simulated by a simulation method. Therefore, the initial phases were optimized by a search of phase interval (from 0 to 2π with a step of 0.05π) of the JFPM method on a public SSVEP dataset using the simulation method with a stimulus duration of 200 ms, resulting in a phase interval of 0.35π between two neighboring frequencies (see Fig. 1b). For more details, please refer to (Wang et al., 2016). The flicker stimulation method uses the sine wave stimulation method, which presents the stimulation by changing the gray-scale of the character flicker background. Thus, when the target they focused on was triggered, P300 and SSVEP were induced simultaneously.

It's worth noting that ISI was varied to explore the impact on BCI performance. In this study, the value of ISI ranged from − 175–0 ms, with an interval of 25 ms. To better evaluate the impact of ISI changes on BCI performance, we defined the radio of the absolute value of ISI to SD as evaluation parameter, which was written as ROIS. Thus, the ROIS values were 87.5%, 75%, 62.5%, 50%, 37.5%, 25%, 12.5%, and 0%, respectively. Taking ROIS equaled to 50% as an example, the overlap time of adjacent stimuli was 100 ms (i.e. ISI=−100 ms). That is to say 100 ms after the last stimulus started flashing, the next stimulus would start flashing. At most two stimuli were flashing at any one time, as shown in Fig. 1c. Similarly, for Fig. 1d, when ROIS was equal to 75%, that was ISI= −150 ms, at most four stimuli were flashing at a certain moment. The stimulation program was developed under MATLAB (MathWorks, Inc.) using the Psychophysics Toolbox Version3.

2.3. BCI Experiment

The subjects were asked to sit calmly in front of the screen with their eyes 60 cm away from the center of the screen. The specific experimental process is shown in the Fig. 2. Taking ROIS equaled to 50% (i.e.
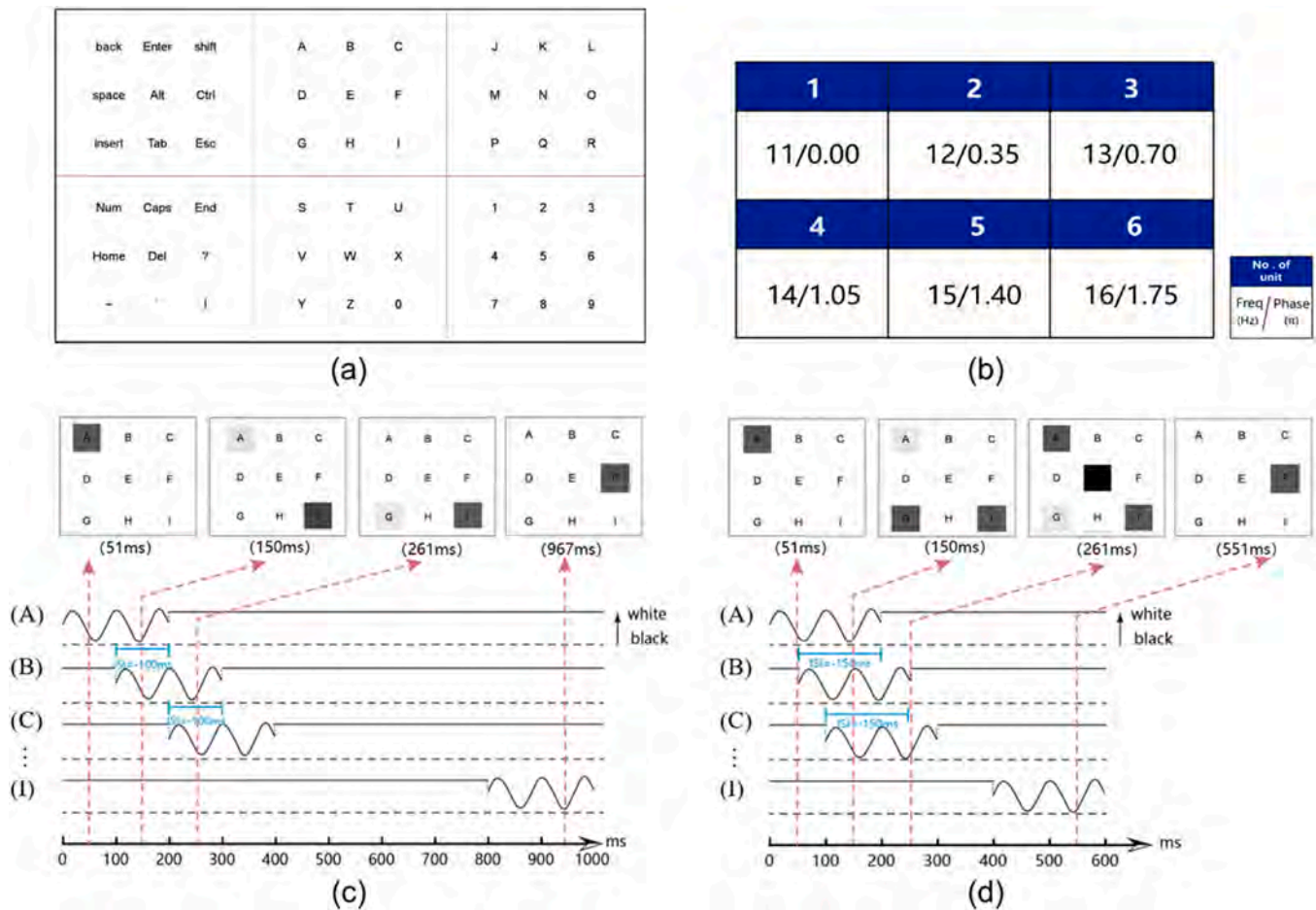
Fig. 1. Illustration of the stimulation in the hybrid BCI speller. (a) Distribution of 54 characters on the screen was divided into 6 sub-spellers by the red line. (b) The selected frequency and initial phase of stimulation squares were displayed for each sub-speller. (c, d) Stimulation process for sub-speller 1, ISI = −100 ms, − 150 ms. Notably, the consuming time of one round was different for ISI= −100 ms and ISI= −150 ms, which was 1000 ms and 600 ms, respectively. The red dotted lines with arrows indicate specific time points.
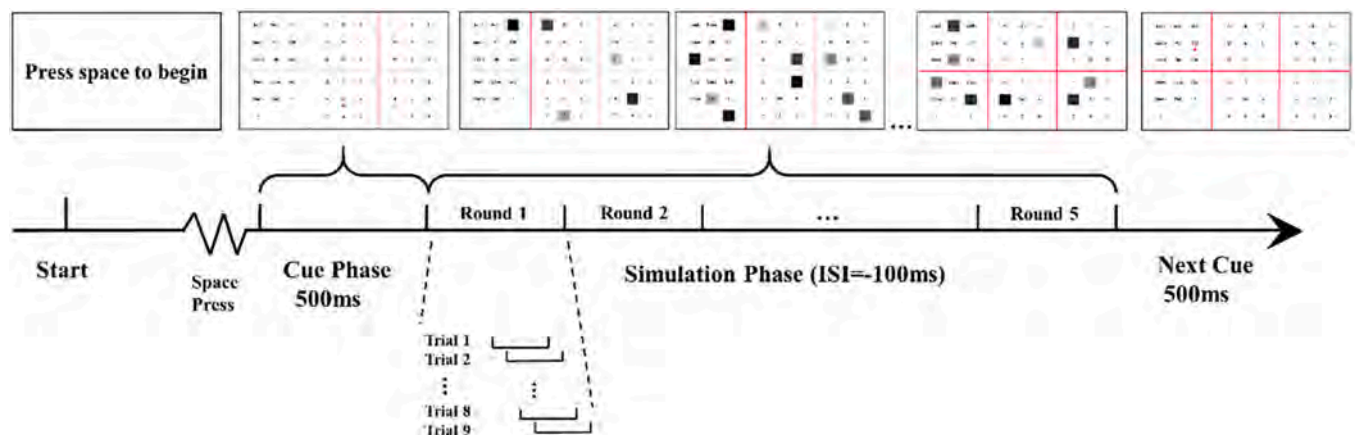
Fig. 2. Experimental flow diagram. The subjects pressed the space key to start the experiment. First the prompt interface appeared for 0.5 s, the stimulation interface began flickering for 5 rounds. After that, the prompt for the next target character would begin.

ISI=−100 ms) as an example, the subjects pressed the space key when they were ready. First of all, a prompt interface were presented on the screen and lasted for 0.5 s. The specified character, which indicated by an underneath read triangle with 1.85 degrees of visual angle for 0.5 s on the prompt interface, was stared by the subjects and also called as the target character. And then the stimulation phase would be began after

that. The visual stimulus ran for five successive rounds for all the characters. Each character flashed for a period of time was called a trial. So, each round consisted of nine trials, of which one was target stimulus, and the other eight were non-target stimuli. The nine characters of each sub-speller were numbered 1–9 from left to right and top to bottom. The sequence of numbers indicated the flashing sequence of the characters.

During the stimulation phase, the subjects were asked to focus on the target character indicated beforehand and count the numbers of times the target was highlighted. All subjects were required to spell all 54 characters for each ISI. In order to reduce the visual fatigue of the subjects, 54 characters under each ISI were randomly and uniformly divided into three sessions. Under each session, the subjects had to spell 18 characters, which were evenly distributed among the six sub-spellers. In the course of the experiment, for each subject, the order of ISI selection was random, and the subjects had to complete 8 groups of ISI experiments. After completing the experiment of each two groups of ISI, the subjects would take a short rest. The whole experiment took about 43 min.

2.4. EEG recording

The EEG data of 64 electrodes were collected by Neuroscan Synamps2 system according to 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 collected by 1000 Hz were stored on the computer after 0.1–200 Hz bandpass filtering and 50 Hz notch filtering.

2.5. EEG processing and feature extraction

As mentioned above, the concurrent P300 and SSVEP features (i.e. hybrid features) are composed of four features, namely frequency-phase-modulated SSVEP and time-modulated P300 features as contained in the traditional hybrid P300 and SSVEP features, and the newly proposed time-modulated SSVEP and frequency-phase-modulated P300 features. According to previous studies (Panicker et al., 2011; Chen et al., 2015b; Xu et al., 2020; Wang et al., 2016), P300 features (i.e. time-modulated and frequency-phase-modulated) are mainly concentrated in the low frequency band ($\leq 10Hz$), while SSVEP features (i.e. frequency-phase-modulated and time-modulated) are distributed in a wider frequency band, and the starting position of the band is determined by the stimulus frequency. Therefore, in this study, stimulation frequencies greater than 10 Hz were adopted so that P300 and SSVEP features could be separated by filters with different bandpass to better analysis.

The signal processing included two parts, i.e. pre-processing and pattern recognition, as shown in Fig. 3. In the pre-processing stage, it mainly includes data filtering, data segmentation, and down-sampling. In the recognition process of the target character, there were two sequential steps: (1) recognizing the sub-speller containing the target character and then (2) recognizing the target character within the identified sub-speller. It's worth noting that P300 could only be used to recognize the target character within the identified sub-speller, while SSVEP and the hybrid features could be used to recognize both the sub-speller containing the target character and the target character within the identified sub-speller. Thus, combined with above features analysis, there were three methods to recognize the target character. The first method was to use only a single SSVEP features. Notably, only when both sub-speller and the target character within the sub-speller are recognized correctly, the output target class is correct, as shown in Fig. 3, otherwise it is incorrect. The second method was to use SSVEP features for sub-speller recognition and P300 features for the target character within the identified sub-speller recognition. In other words, SSVEP and P300 features were combined in a serial manner that was

abbreviated as SSVEP+P300. The third method was to use the hybrid features for the target character recognition. The corresponding specific data processing flow of P300, SSVEP, and hybrid features are introduced as follows, respectively.

For P300, 6 EEG channels (Fz, Cz, Pz, PO7, PO8, Oz) were down-sampled to 200 Hz and filtered by the 1–10 Hz bandpass with Butterworth filters. Then the sampling rate was down to 20 Hz and the P300 feature was extracted from 50 to 800 ms.

For SSVEP, 9 EEG channels (Pz, PO5, PO3, POz, PO4, PO6, O1, Oz, O2) were first down-sampled to 250 Hz and filtered by a filter bank (including seven Chebyshev type I filters) into [X Hz,72 Hz] (X = 9, 17, 25, 33, 41, 49 and 57). SSVEP features were extracted from 140 to 340 ms.

For the hybrid features, 11 EEG channels (Fz, Cz, Pz, PO7, PO3, POZ, PO4, PO8, O1, Oz, and O2) were first down-sample to 250 Hz and filtered by a filter bank (including seven Chebyshev type I filters) into [X Hz, 72 Hz] (X = 2, 9, 17, 25, 33, 41, 49 and 57). The principle of channel selection is based on the feature distribution in brain and fuses more effective features. The hybrid features were extracted from 50 to 450 ms according to our previous study (Xu et al., 2020).

2.6. Step-wise Linear Discriminant Analysis (SWLDA)

SWLDA developed on the basis of Fisher linear discriminant (FLD) analysis is a classic and powerful algorithm for P300, which mainly reduces the space occupied by features by gradually screening and retaining significant features (Krusienski et al., 2006; Xiao et al., 2021). SWLDA is implemented through forward and backward step-by-step analysis. For the input features, the target label category is predicted by using the ordinary least square method to weighted (equivalent to FLD). Since the discriminant function does not contain the initial features, the most statistically significant input features used to predict the target label are added to the discriminant function. When a new feature is entered into the discriminant function, a backward step-by-step analysis is performed to remove the least important feature. The above process is repeated until the discriminant function includes a predetermined number of features or until no new features meet the conditions for adding or removing.

2.7. Ensemble task-related component analysis (TRCA)

The Ensemble TRCA has been proved the most powerful recognition algorithm for SSVEP classification (Nakanishi et al., 2017; Xing et al., 2018). TRCA is an algorithm that finds projection matrix $W = [\omega_{j1} \quad \omega_{j2}... \quad \omega_{N_C}]^T$ to maximize the covariance of task-related components between trial (Tanaka et al., 2013). The h-th trial of EEG signal is $x^{(h)}(t)$. The periods of $x^{(h)}(t)$ are fixed as $t \in [t_h, t_h + T]$, where t_h and T refer to the beginning and the duration of the h-th trial, respectively. The covariance between the h_1-th and h_2-th trials is as follow:

$$C_{h_1 h_2} = \sum_{j_1, j_2 = 1}^{N_C} \omega_{j_1} \omega_{j_2} Cov(x_{j_1}^{(h_1)}(t), x_{j_2}^{(h_2)}(t)) \tag{1}$$

Where N_C is the number of the channels and $Cov(a, b)$ refers to the covariance between a and b. j indicates the indexes of the channels. All possible combinations of trials are summed as:

Fig. 3. The total flow diagram of signal processing.

$$\sum_{\substack{h_1, h_2 = 1 \\ h_1 \neq h_2}}^{N_t} C_{h_1 h_2} = \sum_{\substack{h_1, h_2 = 1 \\ h_1 \neq h_2}}^{N_t} \sum_{j_1, j_2 = 1}^{N_C} \omega_{j_1} \omega_{j_2} Cov\left(x_{j_1}^{(h_1)}(t), x_{j_2}^{(h_2)}(t)\right)$$
$$= \omega^T S \omega$$
$$(2)$$

Where N_t is the number of training trials. Here, the matrix $S = (S_{j_1 j_2})$, $1 \leq j_1, j_2 \leq N_C$ is defined as:

$$S_{j_1 j_2} = \sum_{\substack{h_1 h_2 = 1 \\ h_1 \neq h_2}}^{N_t} Cov(x_{j_1}^{(h_1)}(t), x_{j_2}^{(h_2)}(t)) \tag{3}$$

Considering obtaining a finite result, the following restriction should be met:

$$Var(y(t)) = \sum_{j_1, j_2 = 1}^{N_C} \omega_{j_1} \omega_{j_2} Cov(x_{j_1}(t), x_{j_2}(t)) = W^T Q W = 1 \tag{4}$$

Then, the optimization problem can be transformed as:

$$\widehat{W} = \underset{W}{arg\max}(W^T S W / W^T Q W) \tag{5}$$

The optimal spatial filter is found as the eigenvector of the matrix $Q^{-1}S$ by Lagrange multiplier method. N_f spatial filters corresponding to N_f visual stimulus are integrated as follows:

$$W^m = \begin{bmatrix} W_1^m, & W_2^m, & \ldots\ldots, & W_{Nf}^m \end{bmatrix} \tag{6}$$

Where m is the index of sub-bands designed by filter bank ($m \in [1, N_b]$), N_b is the number of sub-bands. The correlation coefficient between the projection of test data $X(m)$ and averaged individual template $\overline{X}^{(m)}$ is calculated as:

$$r_n^{(m)} = \rho\left(\left(X^{(m)}\right)^T W^{(m)}, \left(\overline{X}_n^{(m)}\right)^T W^{(m)}\right) \tag{7}$$

Where $\rho(a, b)$ indicates the Pearson's correlation analysis between a and b, n indicates the index of sub-spellers ($n \epsilon [1, 6]$). Then correlation coefficients in different sub-bands are weighted by the following equations:

$$\rho_n = \sum_{m=1}^{N_b} \left(m^{-1.25} + 0.25\right) * \left(r_n^{(m)}\right)^2 \tag{8}$$

After that, target can be identified by the following equations:

$$\tau_t = \underset{n}{arg\max} \rho_n \tag{9}$$

2.8. Performance evaluation

In order to evaluate the performance of high-speed BCI, classification accuracy and ITR were used as evaluation indexes in this study, which have been widely used in BCI research (Wolpaw et al., 2000). The ITR can be calculated as follows:

$$ITR = log_2 N + P log_2 P + (1 - P) log_2((1 - P)/(N - 1)) \tag{10}$$

where N is the number of instruction sets, P is the classification accuracy and T is consuming time for each selection. For example, in this study, when ISI $= -150$ ms, the consuming time of T was 1.1 s, 1.7 s, 2.3 s, 2.9 s, and 3.5 s for 1–5 rounds, respectively.

3. Results

3.1. Impact on the hybrid P300-SSVEP BCI performance

As described above, there were three methods for the target

character recognition. The results were as follows. Figs. 4 and 5 showed the averaged accuracies and ITRs of the target character recognition across all subjects using three methods. Three-way repeated measures ANOVA showed that significant main effect of methods (F(2, 22)= 67.75, p < 0.001), ISI conditions (F(7, 77)= 67.49, p < 0.001), and the number of rounds (F(4, 44)= 138.99, p < 0.001) in accuracy. By post-hoc pairwise comparison using paired t-test, the hybrid features achieved the highest accuracy regardless of ISI conditions or the number of rounds, which outperformed than single SSVEP (t_{11} =7.85, p < 0.001) and SSVEP+P300 features (t_{11} =11.48, p < 0.001). The tendency of classification accuracy was consistent to that of ITR. Three-way repeated measures ANOVA also showed that significant main effect of methods (F (2, 22)= 77.92, p < 0.001), ISI conditions (F(7, 77)= 58.96, p < 0.001), and the number of rounds (F(4, 44)= 13.56, p < 0.001) in ITR. Similarly, the hybrid features achieved the highest ITR regardless of ISI conditions or the number of rounds, which was superior to single SSVEP (t_{11} =10.27, p < 0.001) and SSVEP+P300 features (t_{11} =11.95, p < 0.001).

In addition, the effect of ISI on BCI performance using the hybrid features was also analyzed. According to post-hoc pairwise comparison after two-way repeated measures ANOVA (i.e. ISI factor and the number of round factor), the accuracy of ISI= -150 ms or -175 ms were significant lower than that of the other 6 ISI conditions (all p < 0.05). Thus, the ISI= -125 ms was a turning point, after which the accuracy was declined. Notably, because the accuracy increased with the number of rounds, it can still meet the requirements of practical applications. For ITR, the ITR was rising gradually as the ISI decreased. Moreover, there was a significant difference between any pair of 8 ISI conditions (all p < 0.05) except the pair of ISI= -125 ms and ISI $= -150$ ms (t_{11} =2.11, p = 0.06). Therefore, considering accuracy and ITR for hybrid features, the setting of ISI was recommended to set as -175 ms (i.e. ROIS=75%) that achieved the highest ITR of 158.50 bits/min among 5 rounds and all 8 ISI conditions.

3.2. Impact on P300 features

The change of ISI affected the hybrid P300-SSVEP BCI performance. There are many reasons, such as adjacent interference, the amplitude of the induced features as well as data overlap between adjacent trials, and so on. However, it remained unclear how these previous factors would influence the concurrent P300 and SSVEP features. Here, we first analyzed the impact on P300 features including time-modulated P300 and frequency-phase-modulated P300.

Fig. 6 shows the grand average time-modulated ERP at electrode Cz, Pz, and Oz across all subjects and all sub-speller, which were filtered by a band-pass between 1 and 10 Hz. Obviously, P300 was induced between 200 ms and 300 ms after the stimulus onset and amplitudes were different among 8 ISI conditions. To further compare the amplitude of P300, the time window of [0.190 s, 0.325 s] after stimulus onset was selected to calculated for each ISI conditions at Cz and Pz, as shown in Table 1. P300 amplitudes changed slightly until ISI= -150 ms. Two-way repeated measures ANOVA indicated there was a significant difference for ISI factors (F(7, 77)= 8.78, p < 0.001), not for electrode factors (F(1, 11)= 1.05, p = 0.33). According to post-hoc comparison using paired-t test, the P300 amplitude of ISI= -150 ms was significantly lower than that of other ISI conditions (all p < 0.05) except ISI= -75 ms(t_{11} =0.53, p = 0.605), ISI= -100 ms(t_{11} =1.89, p = 0.09), and ISI= -125 ms (t_{11} =1.11, p = 0.29). Moreover, ISI= -175 ms achieved the lowest P300 amplitude, which was significantly inferior than the other 7 ISI conditions (all p < 0.01). The main reason is that the change of ISI affected the target-to-target interval (TTI), which leaded to the change of P300 amplitude(Luck, 2014; Jin et al., 2014).

Fig. 7(a) shows transient ERP of different sub-spellers. The sub-spellers had a variety of P300 at electrode Cz and Pz. Furthermore, P100 at Oz was also different among sub-spellers. Specially, sub-speller
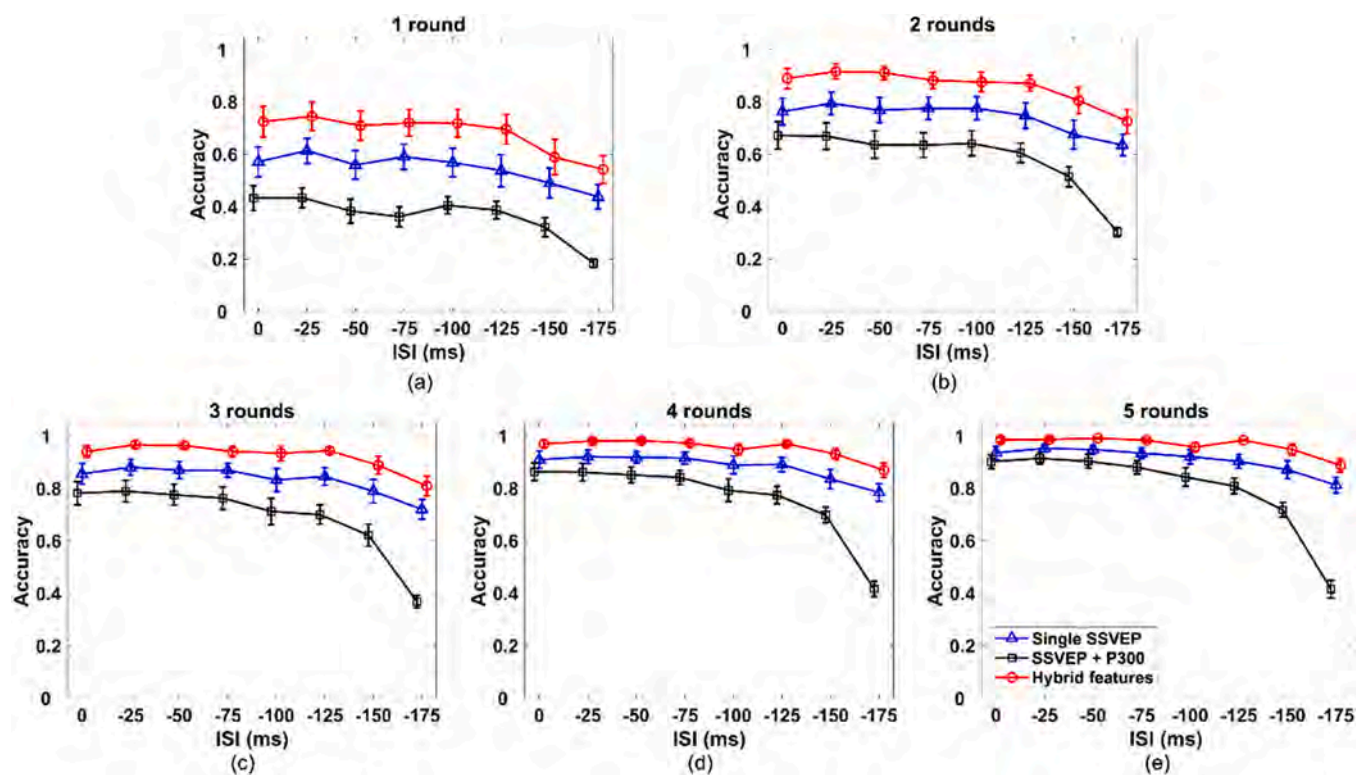
Fig. 4. The averaged accuracy of target character across all subjects were presented from (a) 1 round to (e) 5 rounds with three recognition methods. The error bar indicated standard errors.
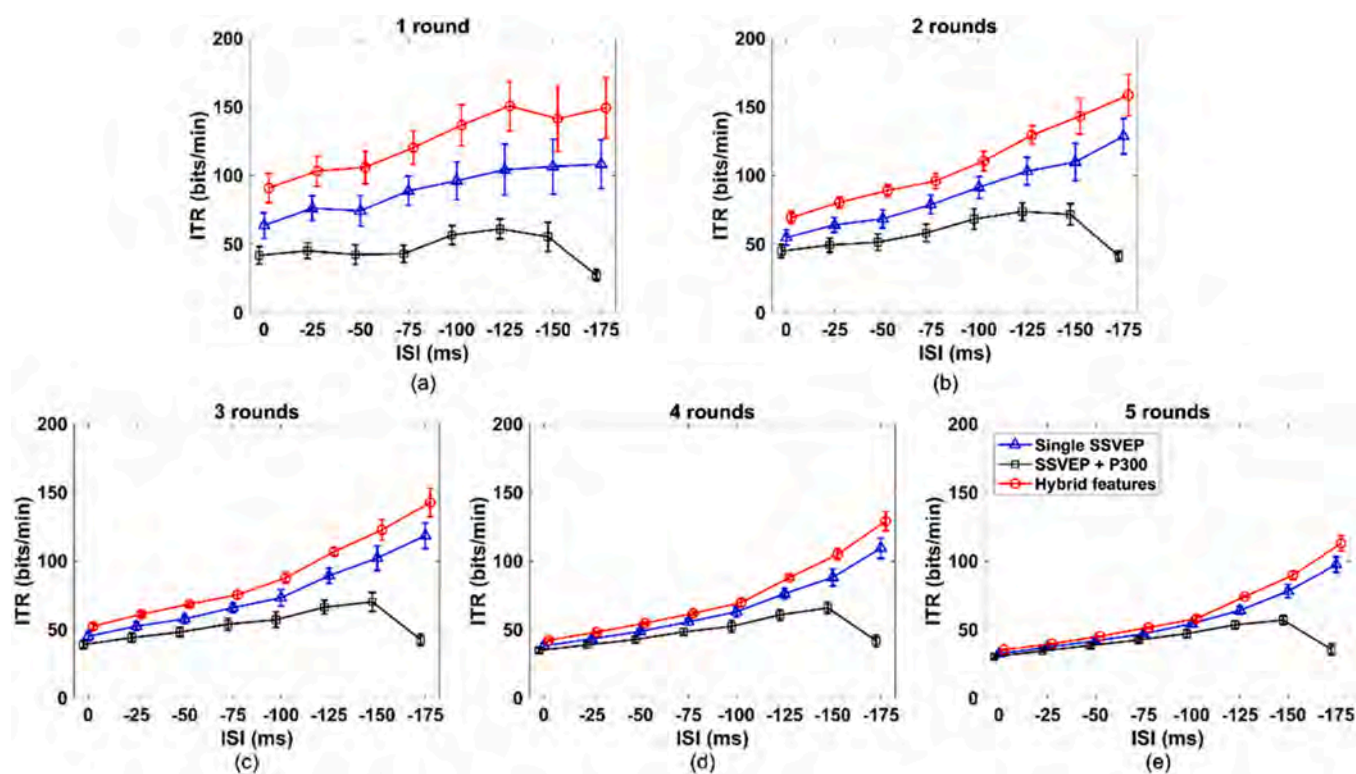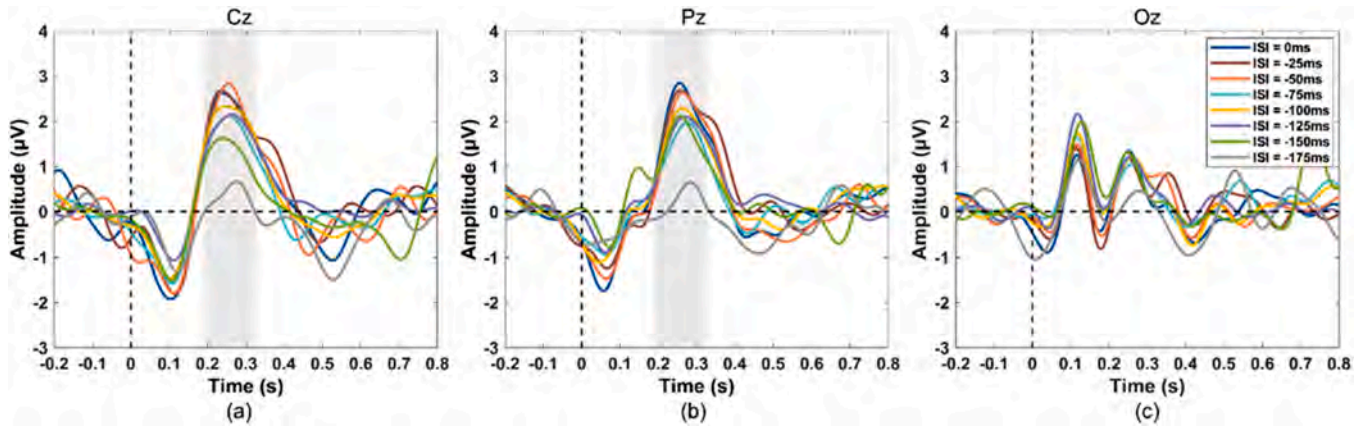

Fig. 5. The averaged ITRs corresponding to accuracy across all subjects were presented from (a) 1 round to (e) 5 rounds with three recognition methods. The error bar indicated standard errors.

Fig. 6. Grand average waveforms of ERP at electrode Cz, Pz, and Oz were presented among eight ISI conditions across all subjects and all sub-spellers. The band-pass filter was set to [1 Hz, 10 Hz] to remove the SSVEP influence. The waveforms were obtained by subtracting the non-target waveforms from the target waveforms. And the baseline of each ISI conditions was corrected according to the data of [− 0.2 s, 0 s] before the onset of stimulation. The gray blocks presented the time window used to calculate P300 amplitude.

Table 1
P300 amplitudes averaged across all subjects.

ISI (ms)	Cz (μV)	Pz (μV)
0	2.21	2.01
-25	2.23	1.97
-50	2.23	1.92
-75	1.78	1.49
-100	2.08	1.70
-125	1.87	1.46
-150	1.34	1.19
-175	0.35	0.27

1 had the largest P300 amplitude while sub-speller 2 had the lowest one. And sub-speller 6 had the largest P100 while sub-speller 3 had the lowest one. The variability of transient ERP might be caused by the non-linear phase resetting of neural oscillations, which is sensitive to the initial conditions of background EEG (Xu et al., 2016; Xu et al., 2016; Xu et al., 2013). Fig. 7(b) showed the transient ERPs of 8 ISI conditions for sub-speller 1. The sub-spellers had different N100 and P300 potentials recorded at Cz and Pz. Moreover, P100 potentials at Oz were different among ISI conditions. Specially, ISI= −50 ms had the largest P300 while ISI= −175 ms had the lowest one. It's consistent with previous study (Luck, 2014). ISI= −25 ms had the largest P100 while ISI= −150 ms had the lowest one. This also indicated that as the ISI decreases, the P300 potentials decrease gradually and other components of ERP also make a difference.
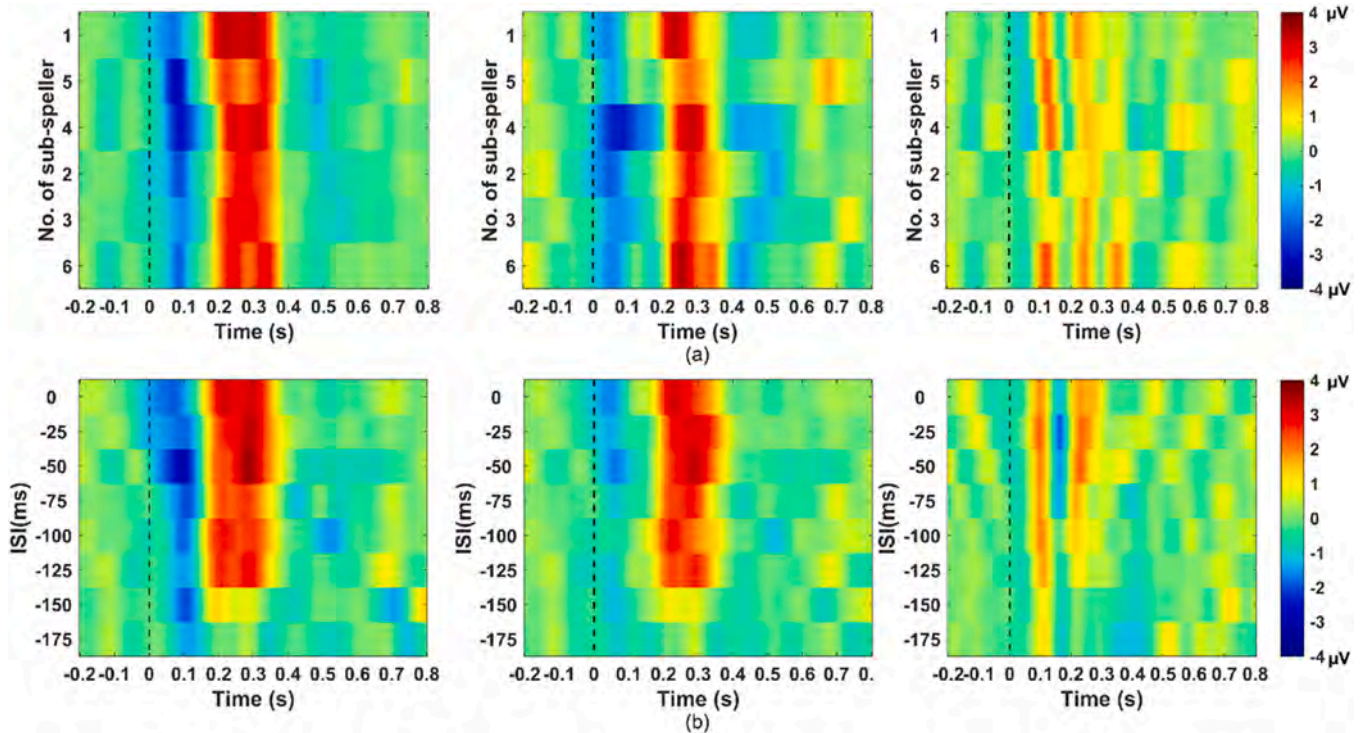
Fig. 7. (a) ERP variations of different sub-spellers for ISI= 0 ms averaged across all subjects were presented at electrode Cz (left), Pz (middle), and Oz (right). The band-pass filter was set to [1 Hz, 10 Hz]. (b) The transient ERPs of different ISI conditions for sub-speller 1 were displayed at electrode Cz (left), Pz (middle), and Oz (right). The band-pass filter was set to [1 Hz, 10 Hz].

3.3. Impact on SSVEP features

In this section, we further analyzed the impact on SSVEP features including time-modulated SSVEP and frequency-phase-modulated SSVEP.

Fig. 8 showed the fundamental SSVEP waveforms at electrode Oz averaged across all subjects. For frequency-phase-modulated SSVEP feature (see Fig. 8a), taking the sub-speller 2 as an example, the SSVEP feature was induced successfully and continued ~200 ms. Moreover, ISI= 0 ms achieved the highest amplitude while ISI= −175 ms achieved the smallest amplitude. It could be seen that the SSVEP amplitudes were decreased with the decrease of the ISI, mainly due to adjacent interference. For example, if ISI was equal to − 175 ms, there were at most 8 characters flashing at a specific moment, whereas there were at most 1 character flashing at a specific moment for ISI= 0 ms. This had a different effect on the induced features. The smaller the ISI value, the greater adjacent interference. In addition, the next cyclic stimulus response for ISI= −175 ms appeared at ~0.5 s (see Fig. 8a), this is because the encoding time of one round is 0.4 s. Similarly, the same phenomenon occurred at ISI= −150 ms that appeared at ~0.7 s

For time-modulated SSVEP, Fig. 8b showed the averaged up envelopes of fundamental SSVEP across all sub-spellers and all subjects. First, the effective SSVEP response lasted ~0.2 s that ranged from ~0.1 s to ~0.3 s, mainly depended on stimulus duration. And then SSVEP amplitudes were gradually decreased with the decrease of ISI value. The reason is the adjacent interference analyzed above. These effects were important reasons for the differences in accuracy among 8 ISI conditions.

4. Discussion

### 4.1. Analyses of the reasons affecting BCI performance

The change of ISI affected the hybrid BCI performance that had many reasons, such as adjacent interference, data overlap between adjacent trials, and so on. In general, adjacent interference refers to the fact that the surrounding stimuli interfere with subject's attention, thereby inducing lower amplitude of features and leading to worse system performance. Some researchers carried out studies on this issue. Jin et al. proposed a new stimulus presentation pattern based on facial expression changes to reduce the adjacent interference, which achieved better performance than control experiment(Jin et al., 2014). In this study, adjacent interference also affected the induced features. Although subjects tried to focus on the target stimulus, the still could not avoid being affected by significant changes in the stimuli adjacent to the target stimulus(Jin et al., 2014). When the ISI value decreased, more

non-target stimuli will be presented at the same time at a specific moment. For example, for ISI= −175 ms, there are a total of 9 stimuli in a sub-speller, one of which is the target stimulus, and the remaining 8 non-target stimuli are presented simultaneously during the experiment, which makes the subjects' attention more likely to be distracted compared to the larger ISI value. As a result, for smaller ISI value, the amplitude of the induced features is lower whether it is SSVEP or P300 features (see Figs. 6, 7, 8).

Another important factor is data overlap between adjacent trials. The degree of data overlap between adjacent trials depended on the ISI value. Theoretically, the characters corresponding to adjacent trials become more and more inseparable as the degree of data overlap increases. Therefore, there are more data overlap between adjacent trials as the ISI decreases, which directly leads to false positives and affects system classification performance. In order to further analyze the impact on system recognition, the confusion matrixes averaged across all subjects of sub-speller 5 for ISI= 0 ms and − 175 ms were selected as examples to explain, as shown in Fig. 9. The flashing pseudo-random sequence for sub-speller 5 was [8 3 7 2 9 5 1 6 4]. For the specified ISI, misclassification was more likely to occur in the category of non-target stimuli that were adjacent to the target stimulus. As shown in the Fig. 9(b), when the category of target stimulus was "5", the predicted category was misclassified to "7" or "9", which accounted for the largest proportion of the misclassified categories. In addition, the decrease of ISI value resulted in an increase in the proportion of misclassification. As shown in the Fig. 9(a)(b), the misclassification rate misidentified from "5" to "9" was 2% for ISI= 0 ms, while the value was increased to 17% for ISI= −175 ms, which occupied in the largest proportion of the misclassified categories for "5" as target. Notably, the non-target stimulus that was adjacent to the target stimulus does not refer to whether it is adjacent in the flashing sequence, but whether the data overlaps between the two trials. For example, if ISI was equal to − 175 ms, there were at most 8 characters flashing at a specific moment, and these corresponding trial data were overlapped.

4.2. Application prospects of optimizing joint time-frequency encoding method

The encoding method of P300-based BCI is time division multiple access (TDMA) that divides multiple targets into multiple time slots. Because of the characteristics of time encoding, an infinite number of commands BCI is theoretically achievable. But the system performance will be extremely slow and cannot meet actual application requirements. The encoding method of SSVEP-based BCI is frequency division multiple access (FDMA) that divides multiple targets into different frequency bands. It supports multiple targets appear simultaneously, so there is the
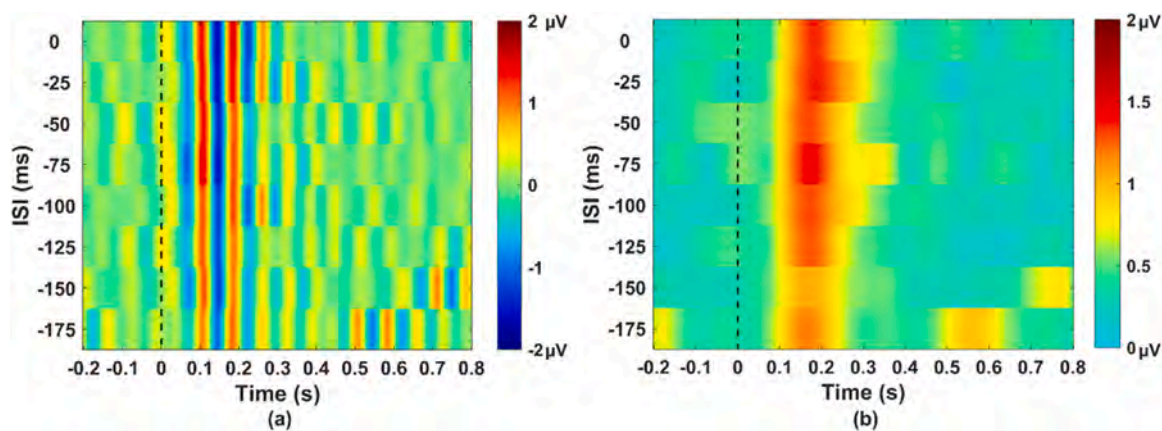
Fig. 8. (a) The averaged fundamental SSVEP waveforms of sub-speller 2 (i.e. 12 Hz) at electrode Oz were presented among 8 ISI conditions across all subjects. (b) Grand average envelopes of fundamental SSVEP components at electrode Oz were displayed across all subjects and all sub-spellers. The envelope was calculated by Hilbert transform. The black vertical indicated stimulus onset. A band-pass filter from 10 Hz to 18 Hz was adopted to remove the influence of the transient ERP.
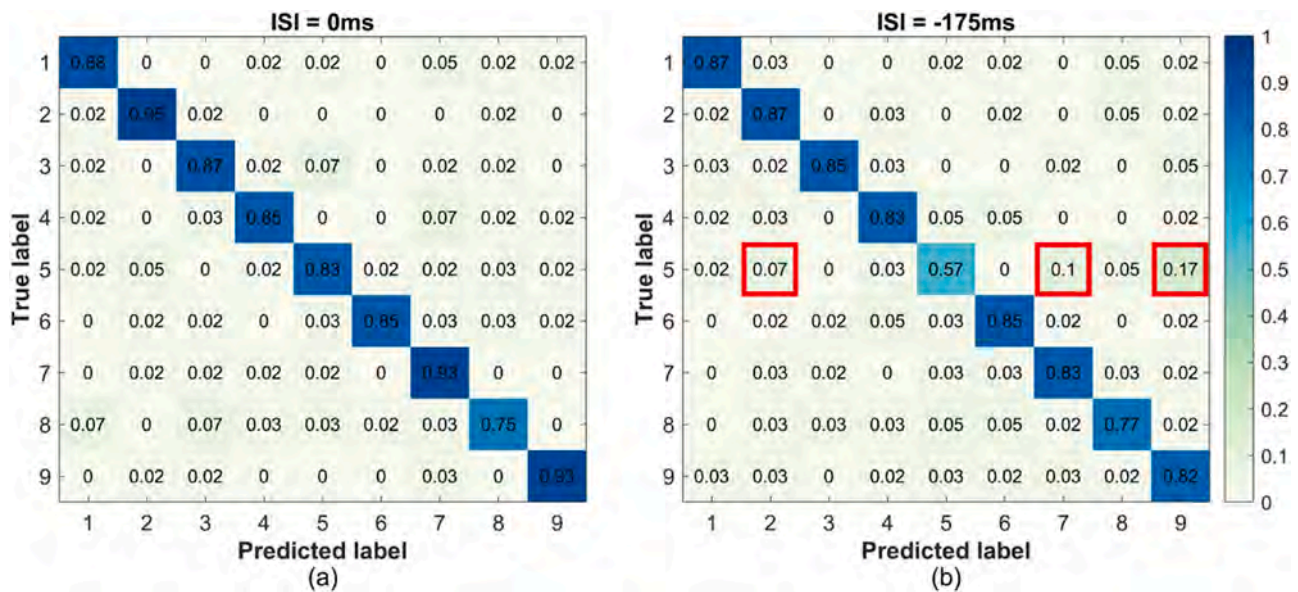
Fig. 9. . (a) and (b) were the averaged confusion matrixes of sub-speller 5 for ISI= 0 ms and ISI= −175 ms across all subjects, respectively.

advantage of fast speed, but the available frequency band is relatively narrow resulting in a limited number of commands. The hybrid BCI used time-frequency joint coding method can combine the advantages of two encoding methods and make up for the shortcomings of two encoding methods, which has been preliminarily proven to achieved better performance (Panicker et al., 2011; Yin et al., 2014; Wang et al., 2015). On this basis, we further proposed the concurrent P300 and SSVEP features and implement a high-speed BCI speller system with over 100 instructions covering full keyboard instructions by using the features. Nevertheless, the hybrid features are needed to be studied, which can further improve BCI performance.

This study explored the effects of ISI on concurrent P300-SSVEP features to find the optimal setting for hybrid BCI. As a result, the hybrid features were achieved the best system performance regardless of ISI or the number of rounds, which was consistent with our previous study (Xu et al., 2020). It indicated that the hybrid features contains more useful information and is suggested to be used in hybrid BCI system in the further. Moreover, it has the advantage of strong robustness and is not susceptible to subject specificity. On the other hand, the results showed that the change of ISI significantly affected system performance. According to the analysis in the result section, the setting of ISI was recommended to set to − 175 ms (i.e. ROIS=75%). Compared to the fixed value of ISI= −100 ms we used in the previous study (Xu et al., 2020), the ITR achieved 149.16 bits/min, 158.50 bits/min, 142.74 bits/min, 129.16 bits/min, and 112.70 bits/min at 1–5 rounds using optimal settings, which have an improvement of 9.15%, 43.31%, 62.57%, 85.10%, and 94.18% at 1–5 rounds, respectively. The marked increases provides a new way to design the high-speed BCI with more than 200 commands.

5. Conclusion

This study explored the effects of ISI on concurrent P300 and SSVEP features and the impact on the BCI performance. The results showed that the concurrent P300 and SSVEP features achieved the highest accuracy and ITRs regardless of ISI or the number of round. And ISI has an nonnegligible impact on hybrid features and system performance. As the ISI decreased, the amplitudes of the induced features and accuracies were decreased gradually, while the ITRs were increased rapidly. It is suggested that ISI can be set to − 175 ms (ROIS=75%), which has an impressive improvement compared to previous study. This study has a guiding significance and a broad application prospect for high-speed BCI

with a large commands.

CRediT authorship contribution statement

Jin Han: Conceptualization, Methodology, Software, Data curation, Writing − original draft, Writing − review & editing. **Chuan Liu, Jiayue Chu**: Software, Validation, Data curation, Formal analysis, Writing − original draft. **Xiaolin Xiao, Long Chen**: Conceptualization, Methodology, Writing − review & editing. **Minpeng Xu, Dong Ming**: Conceptualization, Methodology, Writing − review & editing, Supervision, Project administration.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Allison, B.Z., Pineda, J.A., 2006. Effects of SOA and flash pattern manipulations on ERPs, performance, and preference: implications for a BCI system. Int. J. Psychophysiol. vol. 5 (2), 127–140.

Chen, X., Wang, Y., Gao, S., Jung, T.P., Gao, X., 2015. Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain–computer interface. J. Neural Eng. vol. 12 (4), 046008.

Chen, X., Wang, Y., Nakanishi, M., Gao, X., Jung, T.P., Gao, S., 2015. High-speed spelling with a noninvasive brain–computer interface. Proc. Natl. Acad. Sci. vol. 112 (44), E6058–E6067.

Chen, Y., Yang, C., Chen, X., Wang, Y., Gao, X., 2021. A novel training-free recognition method for SSVEP-based BCIs using dynamic window strategy. J. Neural Eng. vol. 18 (3), 036007.

Chen, Y., Yang, C., Ye, X., Chen, X., Wang, Y., Gao, X., 2021. Implementing a calibration-free SSVEP-based BCI system with 160 targets. J. Neural Eng. vol. 8 (4), 046094.

Feng, J., Yin, E., Jin, J., Saab, R., Daly, I., Wang, X., Hu, D., Cichocki, A., 2018. Towards correlation-based time window selection method for motor imagery BCIs. Neural Netw. vol. 102, 87–95.

Gao, S., Wang, Y., Gao, X., Hong, B., 2014. Visual and auditory brain–computer interfaces. IEEE Trans. Biomed. Eng. vol. 61 (5), 1436–1447.

Jin, J., Allison, B.Z., Sellers, E.W., Brunner, C., Horki, P., Wang, X., Neuper, C., 2011. An adaptive P300-based control system. J. Neural Eng. vol. 8 (3), 036006.

J. Jin, et al. Decreasing the interference of visual-based P300 BCI using facial expression changes. *Proceeding of the 11th World Congress on Intelligent Control and Automation. IEEE*, pp. 2407–2411, 2014.

Jin, J., Daly, I., Zhang, Y., Wang, X., Cichocki, A., 2014. An optimized ERP brain–computer interface based on facial expression changes. J. Neural Eng. vol. 11 (3), 036004.

Krusienski, D.J., Sellers, E.W., Cabestaing, F., Bayoudh, S., McFarland, D.J., Vaughan, T. M., Wolpaw, J.R., 2006. A comparison of classification techniques for the P300 Speller. J. Neural Eng. vol. 3 (4), 299–305.

Luck, S.J., 2014. An introduction to the event-related potential technique. MIT press,, Cambridge, MA.

Martens, S.M.M., Hill, N.J., Farquhar, J., Schölkopf, B., 2009. Overlap and refractory effects in a brain–computer interface speller based on the visual P300 event-related potential. J. Neural Eng. vol. 6 (2), 026003.

Nakanishi, M., Wang, Y., Chen, X., Wang, Y.T., Gao, X., Jung, T.P., 2017. Enhancing detection of SSVEPs for a high-speed brain speller using task-related component analysis. IEEE Trans. Biomed. Eng. vol. 65 (1), 104–112.

Panicker, R.C., Puthusserypady, S., Ying Sun, 2011. An asynchronous P300 BCI with SSVEP-based control state detection. IEEE Trans. Biomed. Eng. vol. 58 (6), 1781–1788.

Polprasert, C., Kukieattikool, P., Demeechai, T., Ritcey, J.A., Siwamogsatham, S., 2013. New stimulation pattern design to improve P300-based matrix speller performance at high flash rate. J. Neural Eng. vol. 10 (3), 036012.

Tanaka, H., Katura, T., Sato, H., 2013. Task-related component analysis for functional neuroimaging and application to near-infrared spectroscopy data. NeuroImage vol. 64, 308–327.

Townsend, G., LaPallo, B.K., Boulay, C.B., Krusienski, D.J., Frye, G.E., Hauser, C.K., Schwartz, N.E., Vaughan, T.M., Wolpaw, J.R., Sellers, E.W., 2010. A novel P300-based brain–computer interface stimulus presentation paradigm: moving beyond rows and columns. Clin. Neurophysiol. vol. 121 (7), 1109–1120.

Townsend, G., Platsko, V., 2016. Pushing the P300-based brain–computer interface beyond 100 bpm: Extending performance guided constraints into the temporal domain. J. Neural Eng. vol. 13 (2), 026024.

Wang, Y., Chen, X., Gao, X., Gao, S., 2016. A benchmark dataset for SSVEP-based brain–computer interfaces. IEEE Trans. Neural Syst. Rehabil. Eng. vol. 25 (10), 1746–1752.

Wang, M., Daly, I., Allison, B.Z., Jin, J., Zhang, Y., Chen, L., Wang, X., 2015. A new hybrid BCI paradigm based on P300 and SSVEP. J. Neurosci. Methods vol. 244, 16–25.

Wang, M., Daly, I., Allison, B.Z., Jin, J., Zhang, Y., Chen, L., Wang, X., 2015. A new hybrid BCI paradigm based on P300 and SSVEP. J. Neurosci. Methods vol. 244, 16–25.

Wang, Y., Gao, X., Hong, B., Jia, C., Gao, S., 2008. Brain-computer interfaces based on visual evoked potentials. IEEE Eng. Med. Biol. Mag. vol. 27 (5), 64–71.

Wang, Y., Wang, R., Gao, X., Hong, B., Gao, S., 2006. A practical VEP-based brain-computer interface. IEEE Trans. Neural Syst. Rehabil. Eng. vol. 14 (2), 234–240.

Wang, K., Xu, M., Wang, Y., Zhang, S., Chen, L., Ming, D., 2020. Enhance decoding of pre-movement EEG patterns for brain–computer interfaces. J. Neural Eng. vol. 17 (1), 016033.

Wolpaw, J.R., Birbaumer, N., Heetderks, W.J., McFarland, D.J., Peckham, P.H., Schalk, G., Donchin, E., Quatrano, L.A., Robinson, C.J., Vaughan, T.M., 2000. Brain-computer interface technology: a review of the first international meeting. IEEE Trans. Rehabil. Eng. vol. 8 (2), 164–173.

Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtscheller, G., Vaughan, T.M., 2002. Brain–computer interfaces for communication and control. Clin. Neurophysiol. vol. 113 (6), 767–791.

Xiao, X., Xu, M., Han, J., Yin, E., Liu, S., Zhang, X., Jung, T.P., Ming, D., 2021. Enhancement for P300-speller classification using multi-window discriminative canonical pattern matching. J. Neural Eng. vol. 18 (4), 046079.

Xiao, X., Xu, M., Jin, J., Wang, Y., Jung, T.P., Ming, D., 2019. Discriminative canonical pattern matching for single-trial classification of ERP components. IEEE Trans. Biomed. Eng. vol. 67 (8), 2266–2275.

Xing, X., Wang, Y., Pei, W., Guo, X., Liu, Z., Wang, F., Ming, G., Zhao, H., Gui, Q., Chen, H., 2018. A high-speed SSVEP-based BCI using dry EEG electrodes. Sci. Rep. vol. 8 (1), 14708.

Xu, M., Chen, L., Zhang, L., Qi, H., Ma, L., Tang, J., Wan, B., Ming, D., 2014. A visual parallel-BCI speller based on the time–frequency coding strategy. J. Neural Eng. vol. 11 (2), 026014.

Xu, M., Han, J., Wang, Y., Jung, T.P., 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. vol. 67 (11), 3073–3082.

Xu, M., He, F., Jung, T.-P., Gu, X., Ming, D., 2021. Current Challenges for the Practical Application of Electroencephalography-Based Brain–Computer Interfaces. Engineering 7 (12), 1710–1712. https://doi.org/10.1016/j.eng.2021.09.011.

Xu, M., Jia, Y., Qi, H., Hu, Y., He, F., Zhao, X., Zhou, P., Zhang, L., Wan, B., Gao, W., Ming, D., 2016. Use of a steady-state baseline to address evoked vs. oscillation models of visual evoked potential origin. Neuroimage vol. 134, 204–212.

Xu, M., Qi, H., Wan, B., Yin, T., Liu, Z., Ming, D., 2013. A hybrid BCI speller paradigm combining P300 potential and the SSVEP blocking feature. J. Neural Eng. vol. 10 (2), 026001.

Xu, M., Wang, Y., Nakanishi, M., Wang, Y.T., Qi, H., Jung, T.P., Ming, D., 2016. Fast detection of covert visuospatial attention using hybrid N2pc and SSVEP features. J. Neural Eng. vol. 13 (6), 066003.

Xu, L., Xu, M., Ma, Z., Wang, K., Jung, T.P., Ming, D., 2021. Enhancing transfer performance across datasets for brain-computer interfaces using a combination of alignment strategies and adaptive batch normalization. J. Neural Eng. vol. 18 (4), 0460e5.

Xu, M., et al., 2018. A brain–computer interface based on miniature- event-related potentials induced by very small lateral visual stimuli. IEEE Trans. Biomed. Eng. vol. 65 (5), 1166–1175.

Yin, E., Zhou, Z., Jiang, J., Chen, F., Liu, Y., Hu, D., 2014. A speedy hybrid BCI spelling approach combining P300 and SSVEP. IEEE Trans. Biomed. Eng. vol. 61 (2), 473–483.