IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 6, 2020, accepted January 4, 2021, date of publication January 11, 2021, date of current version January 20, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3050545

The Effect of SOA on An Asynchronous ERP and VEP-Based BCI

MENGFAN LI[1,2], (Member, IEEE), GUANG YANG[1,2], ZHEN LIU[1,2], MINGHONG GONG[1,2], GUIZHI XU[1,2], (Member, IEEE), AND FANG LIN[3]

[1]State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of Technology, Tianjin 300401, China
[2]Tianjin Key Laboratory of Bioelectromagnetic Technology and Intelligent Health, Hebei University of Technology, Tianjin 300401, China
[3]Neuracle Technology Company Ltd., Beijing 100097, China

Corresponding author: Mengfan Li (mfli@hebut.edu.cn)

This work was supported in part by the Natural Science Foundation of Hebei Province under Grant F2018202088, and in part by the National Natural Science Foundation of China under Grant 61806070, Grant 51977060, and Grant 51737003.

ABSTRACT Our previous study established an asynchronous BCI system by using the oddball paradigm to simultaneously induce event-related potentials (ERPs) and visual evoked potentials (VEPs) (E-V BCIs). We found that stimulus onset asynchrony (SOA) is an important factor for performance since it significantly affects the ERP and VEP. Increasing the SOA increases the ERP, which improves the accuracy of detecting target stimuli. However, a larger SOA leads to a lower VEP frequency, which causes the VEP to have poor accuracy when discriminating between the brain states. How to balance the two potentials and accuracies is a problem. This study established eight SOAs from 100 ms to 375 ms that were composed of different interstimulus intervals and the same stimulus duration of 80 ms. We used a probability-based Fisher linear discriminant analysis (P-FLDA) classifier to calculate the classification accuracies of the ERP-based visual speller, VEP-based brain state discrimination, and E-V BCI. The results show that as the SOA increases, the amplitudes of N200 and P300 increase, and the accuracy also shows an increasing trend. However, the frequency of the VEP and the accuracies of state discrimination show downward trends. The change in accuracies of the E-V BCI system combining these two parts is nonlinear, and the SOA optimal value is 125 ms. The SOA of 125 ms yields the best accuracies of 95.83% and practical bit rate of 57.17 bits/min in the E-V BCI system, which provides a guideline for selecting the SOA to improve the performance.

INDEX TERMS Brain–computer interfaces, asynchronous system, stimulus onset asynchrony, event-related potential, visual evoked potential.

I. INTRODUCTION

Brain-computer interface (BCI) is a technology that decodes brain signals to control instructions to an external device between a human and a computer [1]–[4]. BCI has been applied in many fields, such as medicinal areas [5], robotic arms [6] and communication [7]. Event-related potentials (ERPs), e.g., the N200 and P300 potentials, are brain potentials representing the peak of the cerebral cortex in the fixed latency period after the occurrence of low-probability events [8]. Visual evoked potentials include steady-state visual evoked potentials (SSVEPs) and transient visual evoked potentials (TSVEPs). SSVEP and TSVEP refer to the potentials corresponding to the stimulation frequencies after human vision is stimulated by more than 6 Hz and less than 6 Hz, respectively [9]. The ERP and VEP are widely used in BCI, such as for wheelchair control [10] and rehabilitation [11].

Robots and vehicles are commonly used in BCI because of their great contributions to industries and rehabilitation [12], [13]. Visual stimuli could improve the performance of an ERP-based BCI system by increasing potential amplitudes. Our previous study compared the effect of images filled with a pure background and humanoid robot images on an ERP-based BCI. The results show that humanoid robots induce larger N200 potentials in the frontal and central areas [14], which improves BCI performance. Robot and vehicle stimuli also use familiarity as human face stimuli to increase amplitudes and improve accuracy [15]. Therefore, we use robotic arms and vehicles as visual stimuli to

The associate editor coordinating the review of this manuscript and approving it for publication was Larbi Boubchir.

induce ERP. BCIs based on visual stimuli include synchronous and asynchronous systems. Studies on asynchronous systems have received increasing attention since they permit the subject to control devices at his/her pace. The system only outputs instructions when the subject's brain is in the working state and does not process them when the subject is idle so that the subject can control the device flexibly and freely according to their own wishes. The challenge of an ERP-based asynchronous BCI is to discriminate subjects' working and idle states [16]. Many researchers have combined the ERP with SSVEP and motor imagery to recognize states by using the non-ERP paradigm as the state switch [17]–[19]. Our previous study proposed using only one odd-ball paradigm to realize asynchronization by simultaneously inducing the ERP and TSVEP and combining their classifications [20], [21].

In our proposed BCI system combined with ERP and VEP (E-V BCI), we found that the setting of the stimulus onset asynchrony (SOA) was very important since it could influence the shape of the ERP and VEP. The SOA is the time between the onsets of two consecutive stimuli [22]. It represents the number of stimuli appearing in a fixed time; in other words, it stands for the probability of a stimulus appearing in the time aspect. The probability would affect the ERP since the ERP is related to the renewal of working memory. Don-chin reported that a longer SOA is more conducive to memory renewal and P300 formation [23]. Brendan *et al.* found that the amplitudes of ERP decreased with shorter SOAs [24]. An interval that is too short will lead to problems such as overlap between signals and repetition blindness, which makes the ERP amplitudes small [25]. Therefore, increasing the SOA is good for the E-V BCI since a higher ERP can improve the classification accuracy of detecting target stimuli. In contrast, improving the classification accuracy of discriminating the work and idle states needs to decrease the SOA. The formation of the VEP is based on a constant switch between the onset and offset of all stimuli, so the frequency of the VEP is closely related to the SOA. Hu *et al*. believed that a lower frequency of stimulation could lead to fatigue in the subjects and degradation of the system performance [26]. Daisuke also found that a high frequency of VEP reduced the probability of epileptic induction [27]. Our previous findings match with these researchers' works: increasing the SOA leads to an increasing trend of ERP recognition accuracy and a downward trend of VEP-based recognitions accuracy. The contradiction between the ERP-based command recognition and VEP-based working state recognition poses a great challenge to the asynchrony E-V BCI when setting a SOA to yield both high accuracy and speed. Hohne *et al.* found 175ms most suitable for the synchrony ERP BCI system [28]. Whether it is also best for an asynchrony ERP BCI system; and which one is more suitable if 175ms is not the best, are studied in this current work to obtain the best performance of E-V BCI system.

To find the optimal SOA value, this study sets up eight SOAs including the same stimulus duration (STD) and different inter-stimulus time (ISI) to conduct E-V BCI experiments. The result show that the SOA optimal value is 125 ms that yields the highest accuracies of 95.83% and practical bit rate (PBR) of 57.17 bits/min in the E-V BCI system. To our best of knowledge, we are the first to find the optimal SOA value for an asynchronous system. The contribution of this study is to give a guideline of selecting the SOA to improve the performance of asynchronous systems. For an asynchronous ERP BCI system, 125 ms is the best, instead of 175 ms that is widely used in the synchronous system, indicating that the optimal SOA should be different based on the system type.

II. MATERIALS AND METHODS

A. PARTICIPANTS AND DATA ACQUISTION

Nine healthy subjects (23-26 years old, 24.8±2.37) with normal or corrected-to-normal vision took part in this experiment. All subjects were right-handed and had no history of psychiatric disorders. They all signed voluntary consent before the experiment and were told all possible consequences of the study and the experimental procedure. The experiment in this study was approved by the Ethics Review Committee of Hebei University of Technology (Number: HEBOThMEC2019001).

The study used the Neuroscan Synamps2 system to collect 14 channels (Fz, FCz, Cz, CPz, CP3, CP4, Pz, P3, P4, T5, T6, Oz, O1, O2) of electroencephalogram (EEG) signals at a sampling rate of 1000 Hz [29]. These 14 channels locate in the sagittal line and posterior area of the brain that is helpful to induce high ERP and VEP and reach better performance of the BCI system [30], [31]. The impedance of the electrodes was 5kΩ [32]. Reference electrodes were placed at the bilateral mastoids. AFz was the ground.

B. EXPERIMENTAL PARADIGM DESIGN

In the E-V BCI system, the state in which the subject watches an odd-based visual speller to choose a stimulus is a working state; the state that subject watches the video shot to get known of the environment and needn't to output command is an idle state.

1) VISUAL SPELLER

The visual speller is a 4×4 pattern composed of 16 command images. The command images indicate either a motion performed by a robotic arm or a vehicle. Table 1 lists the content of these images. ''v'', ''ra'' and ''vc'' represent the vehicle, robotic arm and vehicle camera, respectively. The resolution and size of the command image are 794 × 596 pixels and 4.12 × 3.41 cm, respectively. The interface adopts the row and column method [33] to flash the command images. Each row/column of command images flashes randomly. When a command image finishes flashing, it is replaced by a shield image. The shield image consists of a black square and a white point. Figure 1(a) and (b) show that the third column and the third row flash, and the others are shielded by shield

TABLE 1. Visual stimulus.

	Column 1	Column 2	Column 3	Column 4
Row 1	$Forward_{(v)}$	$Retreat_{(v)}$	$Turn\ Left_{(v)}$	$Turn\ Right_{(v)}$
Row 2	$Forward_{(ra)}$	$Retreat_{(ra)}$	$Turn\ Left_{(ra)}$	$Turn\ Left_{(ra)}$
Row 3	Arm Open	Arm Close	$Forward_{(vc)}$	$Retreat_{(vc)}$
Row 4	$Turn\ Right_{(vc)}$	$Turn\ Left_{(vc)}$	Ultrasonic Sensor	Infrared Sensor

images. The command image that the subject focuses on is the target stimulus, and the subject ignores is the nontarget stimulus. The process in which one row/column is shown and shielded is a flash. A repetition is that all rows and columns flash once and contains 12 flashes, and 10 repetitions constitute a trial.

The SOA is the stimulus interval in the visual speller, which is from the last command image to the next command image [24]. It includes the STD in which the stimuli are shown and the ISI, which is the interval between two stimuli shown. The SOA ranges from 100 ms to 375 ms, in which the ISI ranges from 20 ms to 295 ms and the STD remains at 80 ms. The 100 ms and 375 ms are set to guarantee high amplitude of ERP and PBR of the system. In order to take 150 ms, 175 ms and 225 ms that are widely used in BCI into experiment design, we set 25ms interval from 100 ms to 175 ms, and 50 ms interval from 175 ms to 375 ms [34], [28], [35]. Each SOA value corresponds to a flashing frequency. The frequency (Hz) is 1/SOA (ms)$\times$1000. Table 2 shows the eight SOAs and the corresponding frequencies.

TABLE 2. The value of SOA in the experiment.

	SOA(ms)	STD(ms)	ISI(ms)	Flashing frequency(Hz)
1	100	80	20	10.00
2	125	80	45	8.00
3	150	80	70	6.66
4	175	80	95	5.71
5	225	80	145	4.44
6	275	80	195	3.64
7	325	80	245	3.08
8	375	80	295	2.66

2) VIDEO WATCHING

This study puts the subject in an idle state by letting him watch a video. The subjects are in a relaxed state and have no need to output a command. The video is recorded by a camera mounted on the platform of the vehicle with a resolution of 640 $\times$ 480 pixels, as shown in Figure 1 (c). The video plays the environment around the vehicle in the first-person view. The driving environment is a laboratory. This video watching experiment simulates daily situations of the BCI task: the subject watches the video window to become familiar with the environment or make a decision without intending to output a command.

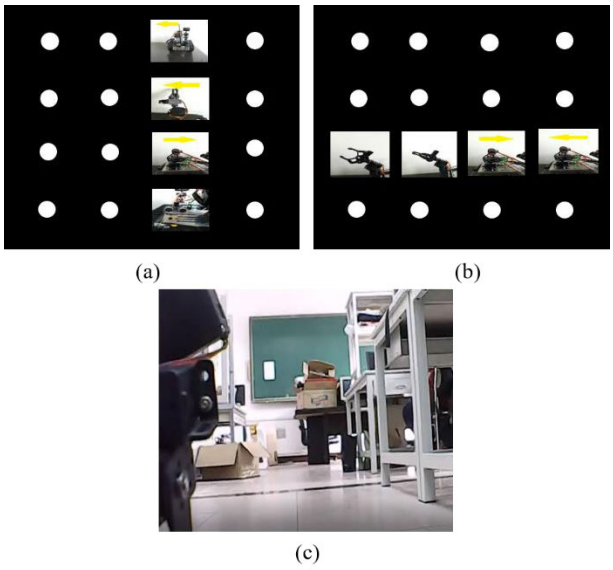

FIGURE 1. Visual speller (a) The third-column command images flashes. (b) The third-row command images flashes. (c) Video of the vehicle.

C. EXPERIMENTAL PROCEDUCE

Subjects are seated 60 cm in front of the two 19-inch LCD computer screens. The screens both have a resolution of 1920 $\times$ 1080 pixels and a screen refresh rate of 60 Hz. The two screens separately display the visual speller (left screen) and the video (right screen).

In eight E-V BCI experiments, all the settings are the same except for the SOAs. Figure 2 (a) shows the components of the eight SOAs. Figure 2 (b) is the flow chart of the E-V BCI system experiment. Each subject is required to perform eight E-V BCI experiments in random order of SOA with a 2-minute rest between experiments [36]. In each E-V BCI experiment, the brain is in the working state when the subject performs the visual speller and the idle state when the subject is watching the video. The subjects' brain states switch between working and idle states. The subject first stares at the target stimulus in the visual speller to output an instruction for a trial time and ignores video watching. Then, the subject switches his attention from the visual speller to the video watching, and the interval time is set to 500 ms, which is the shortest interval for the subjects to change their perspective [37]. During this period, the system does not recognize the EEG. The subject focuses on video watching to observe environments in front of the vehicle for a trial time. When watching the video, the subject looks at the content of the video on the right screen and does nothing while the left screen flashes the stimuli as usual. The subjects completed 16 trials of visual speller and 16 trials of video watching in one E-V BCI experiment. Each subject is required to complete 8 E-V BCI experiments with different SOAs.

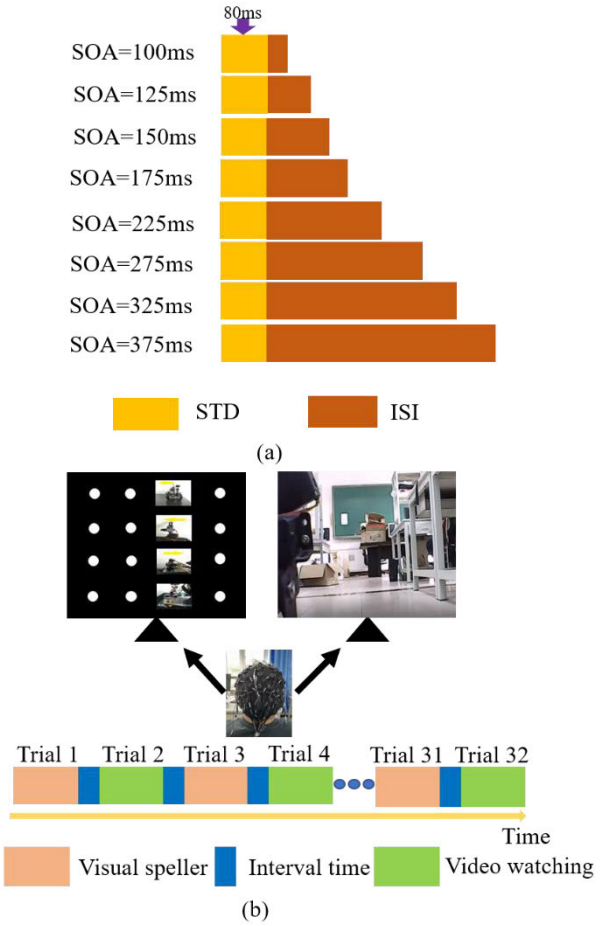

FIGURE 2. The experimental process.

D. FEATURE EXTRACTION AND CLASSIFICATION

The feature extraction section describes how to extract time features of the ERP, and how to extract the frequency features of the TSVEP.

1) ERP FEATURES

For each visual stimulus, firstly, the EEG data from 50 ms to 750 ms poststimulus are cut. Secondly, a 3rd-order Butterworth filter (0.1-20 Hz) is used to filter the data segment to reduce power frequency interference and other noise [38]. Then, the 50 ms data prestimulus is used as the baseline for correction to reduce the slow potential drift and other artifacts. Finally, the data are downsampled to 40 Hz to obtain 28-dimensional feature vector from each channel [39], [40]. This paper adopts 14 channels as the feature channels to yield a 392-dimensional feature. Artifact removal is not performed since the ERP is obvious [14], [41].

2) TSVEP FEATURES

The data from 1000 ms prestimulus to 1000 ms poststimulus are cut for each flash. Each epoch corresponds to a flash; therefore, a repetition yields eight epochs. Then, the eight epochs in a repetition are averaged to obtain a data segment. We use a 3rd-order Butterworth filter to filter every segment from 0.1 Hz to 20 Hz. Finally, we use a fast Fourier transform (FFT) to extract the amplitude characteristics in the frequency domain of each segment. FFT is a common method used to extract signals in the frequency domain of the VEP. It is a fast method of FT to shorten the operation time [42], [43].

$$x(k) = \sum_{N-1}^{n=0} x(n) \times e^{\frac{-j2\pi kn}{N}} \quad (k = 0, 1 \ldots, N-1) \qquad (1)$$

where $\mathbf{x}(n)$ is the EEG signal at time point n, N is the length of a data epoch, and k is the frequency.

The feature extraction part of the algorithm selects the data from 0 Hz to 20 Hz to form a 20-dimensional feature vector. The VEP potential is usually in the occipital region of the brain [44], so the feature channel is Oz.

3) CLASSIFICATION

This study uses the P-FLDA [21] to distinguish the subject's brain state and the target stimulus. P-FLDA is a method which transforms the value calculated by the FLDA into a probability. It has the advantage of being able to translate the results of different data into the same range, narrowing the order of magnitude between the data. ERP and VEP are fused through the P-FLDA classifier to form the probability value, and the result of the fusion is used to judge the idle state and target stimulus of the brain.

E. EVALUATION

This study uses accuracy and PBR to evaluate the performances of the E-V BCI and E BCI.

1) ACCURACY

The equation for calculating accuracy is as follows:

$$Acc = \frac{N_c}{N_{ALL}} \qquad (2)$$

where Nc is the number of trials judged correctly, and N_{ALL} is the total number of trials.

This study included three types of accuracy: $\mathrm{Acc}_{\text{visual speller}}$, $\mathrm{Acc}_{\text{videowatching}}$ and $\mathrm{Acc}_{\text{E-V BCI}}$. The $\mathrm{Acc}_{\text{visual speller}}$ value represents the ratio between the number of trials in which the target stimuli are judged correctly to the total number of trials under the visual speller condition. The $\mathrm{Acc}_{\text{visual speller}}$ indicates the accuracy of sending instructions correctly, which only uses the ERP features to detect the target stimulus. The $\mathrm{Acc}_{\text{video watching}}$ value represents the ratio between the number of trials in which the idle states and the working state are judged correctly and the total number of trials under the visual speller and video watching experiments. $\mathrm{Acc}_{\text{videowatching}}$ indicates the accuracy in recognizing brain states, which only uses the VEP features to discriminate the states. The $\mathrm{Acc}_{\text{E-V BCI}}$ value represents the ratio between the number of trials in which the target stimuli or the idle

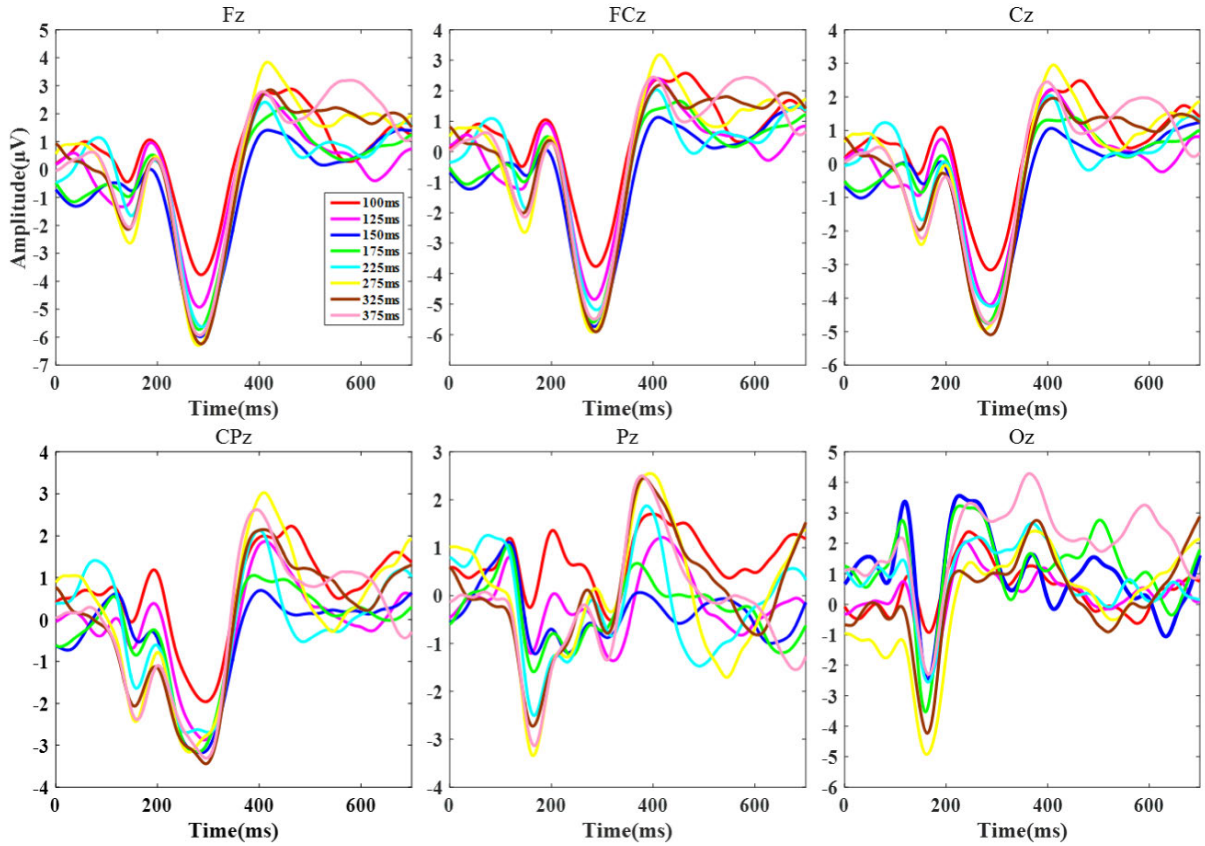

FIGURE 3. The waveforms of target stimulus in Fz, FCz, Cz, CPz, Pz, and Oz channels.

states are judged correctly and the total number of trials in both the visual speller and video watching experiments. The $\mathrm{Acc}_{\mathrm{E-V\,BCI}}$ indicates the accuracy in recognizing brain states and sending instructions correctly, which uses both the ERP features and the VEP features to detect the two states and the target stimulus.

2) PBR

PBR is used to estimate the speed of the system in a real-world setting [45]. The PBR is calculated using $\mathrm{Acc}_{\mathrm{E-V\,BCI}}$ and information transfer rate (ITR). The equation is as follows:

$$PBR = ITR \times [1 - 2 \times (1 - Acc)] \tag{3}$$

$$ITR = \left\{ lbQ + Acc + lbAcc + (1-Acc) \times lb \times \frac{(1-Acc)}{Q-1} \right\} \times M$$

$$M = \frac{60 \times 1000}{T \times N_r + t} \tag{4}$$

where Q is the number of visual stimuli, M is the number of commands output in one minute, N_r is the number of repetitions in a trial, T is the duration of a repetition, and t is the time to switch between visual speller and video watching, which is set as 500 ms.

3) STATISTICAL ANALYSIS

This paper uses the Wilcoxon signed-rank test to investigate the differences between the different SOA conditions. The significance level is set at 0.05 [46].

III. RESULT

This section shows the results of the experiment. In addition, all the results are averaged by nine subjects. The ERP amplitude, classification accuracy and PBR obtained constitute significant differences in statistical analysis ($p < 5 \times 10^{-2}$).

A. TIME-DOMAIN ANALYSIS

Figure 3 shows the waveforms of N200 and P300 in Fz, FCz, Cz, CPz, Pz, and Oz induced by the target stimulus. The time range is from 0 ms to 700 ms. In the frontal and central areas (Fz, FCz, Cz and CPz), N200 and P300 are evoked at approximately 290 ms and 410 ms after stimulus. In the parietal and occipital areas, N200 potential and P300 potential are induced at approximately 190 ms and 390 ms after stimulus. The amplitudes of N200 and P300 increase gradually as the SOA increases. Table 3 shows the P100 peak to N200 peak and the N200 peak to P300 peak. we can reference out the variability caused earlier in the waveform and gain reliable insights concerning the peaks of N200 and P300. The P100 peak to N200 peak reaches the highest and the lowest amplitudes

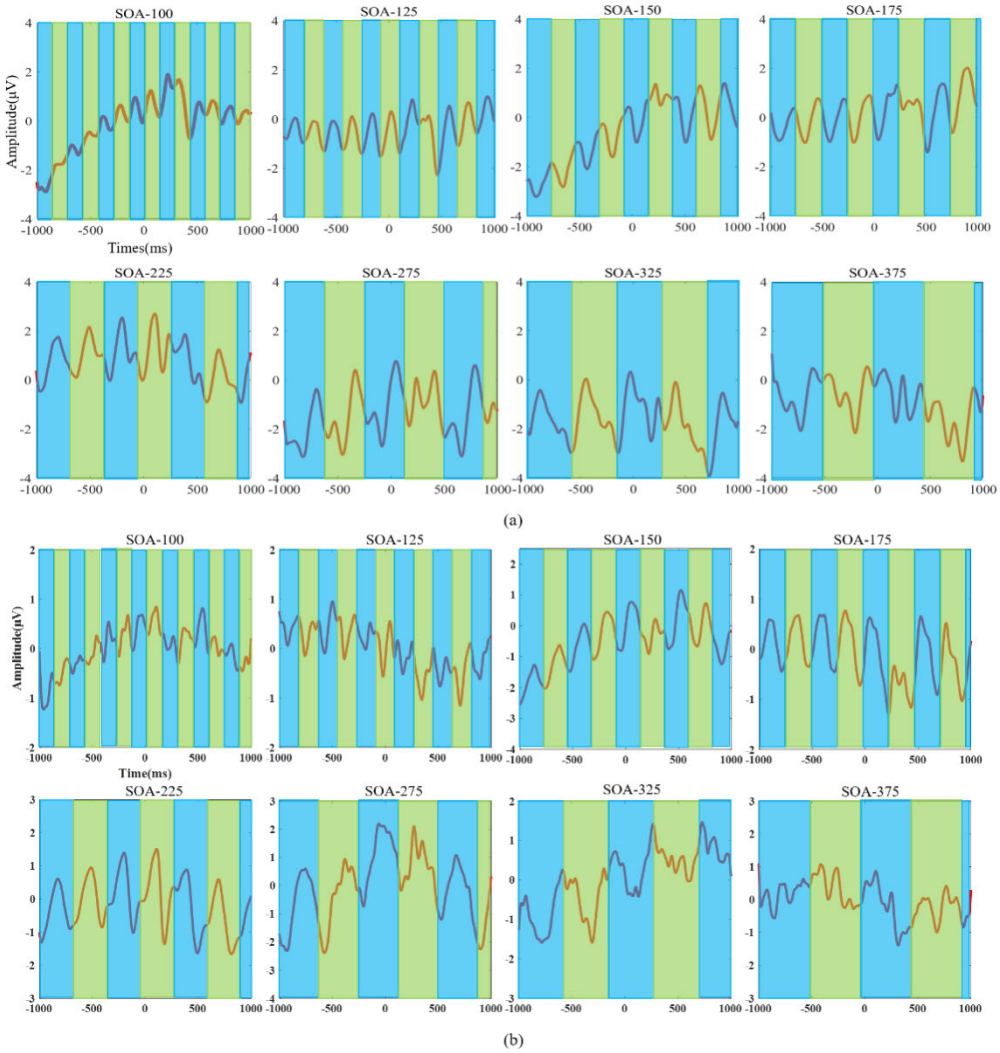

FIGURE 4. The waveforms of nontarget stimulus in Fz channel. (b) The waveforms of nontarget stimulus in Oz channel.

of 5.6 μV and 4.1 μV when the SOA was 125 ms and 150 ms, respectively, and the N200 peak to the P300 peak reaches the highest and the lowest amplitudes when the SOA was 275 ms and 150 ms, 8.0 μV and 5.2 μV, respectively.

To show the periodic change in waveform induced by the visual stimulus clearly, we plot the waveforms evoked by the nontarget stimulus in Figure 4. The colored squares in Figure 4 mark these waveforms with a cycle in channel Fz and Oz. These two channels have the same trend: As SOA increases, the waveform cycle changes. This tendency of becoming dense is caused by the increase of the SOA. Attention can cause changes in the FZ channel [47], and the OZ is the main channel that produces ERP.

B. FREQUENCY-DOMAIN ANALYSIS

Figure 5 shows the frequency spectrum of signal under the eight visual speller conditions and video watching condition.

TABLE 3. The amplitude OF N200, P300 IN Cz, AND VEP IN Oz.

SOA(ms)	P100 peak to N200 peak(μV)	N200 peak to P300 peak(μV)	VEP(μV)	Frequency (Hz)
100	4.4	5.3	0.09	6.71
125	5.6	6.5	0.19	5.49
150	4.1	5.2	0.26	4.27
175	5.0	6.0	0.26	4.02
225	4.5	6.2	0.28	3.17
275	5.0	8.0	0.28	2.56
325	5.3	7.2	0.24	2.31
375	4.8	7.1	0.28	2.07

The frequency range is from 0 Hz to 20 Hz. The vertical coordinate is the amplitude. When the subject is focusing on the visual speller, the peak spectrum from 6.71 Hz to 2.07 Hz ($p = 4.7 \times 10^{-2}$) as the SOA increases, as shown in Table 3. These peak spectra are mainly concentrated in 0 Hz to10 Hz.

Compared with the peak spectrum induced by the visual speller, that of video watching is very small: 0.03 μV

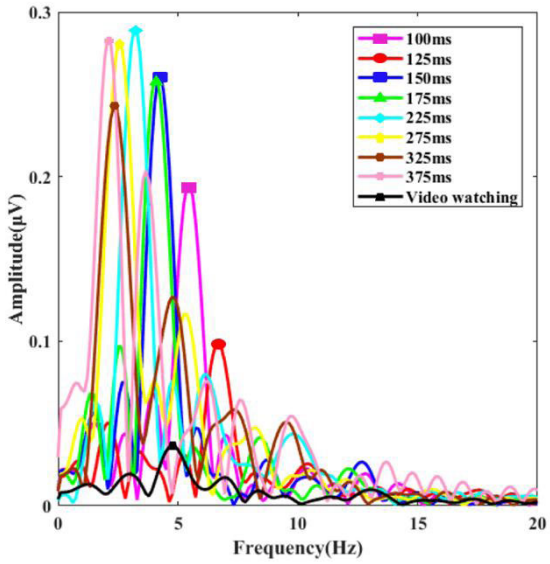

FIGURE 5. The frequency domain waveforms of VEP in Oz channel.

at 4.88 Hz. In the frequency domain, the amplitude peak differences induced by the visual speller and video watching provide a good basis for distinguishing the two types of signals.

C. THE ACCURACY ANALYSIS

Figure 6 plots $Acc_{visual\ speller}$, $Acc_{videowatching}$ and $Acc_{E-V\ BCI}$ across the 9 subjects. With the increase of repetition, the three kinds of accuracy all show an increasing trend. While, the three accuracies vary differently as SOA increases. When the SOA increases, the $Acc_{visualspeller}$ increases from 91.67% to 95.83%, while the $Acc_{video\ watching}$ decreases from 98.96% to 90.97%. The opposite trends of $Acc_{visual\ speller}$ and $Acc_{video\ watching}$ leads to a non-linear variation in $Acc_{E-V\ BCI}$. When SOA is 125 ms, the $Acc_{E-V\ BCI}$ is the highest of 95.83% and the corresponding PBR also reaches the highest of 57.17 bits/min. The results indicate that the optimal SOA value for an asynchronous BCI system is 125 ms.

D. THE PBR OF E-V BCI

Figure 7 shows the PBR of the E-V BCI system under eight SOAs. Firstly, with increasing repetition, the PBR increases and then decreases. And the PBR reaches the highest when the repetitions are two. Secondly, as the SOA increasing, the PBR tends to decline. When the SOA is 125 ms, the PBR reaches the highest of 57.17 bits/min. This indicates that PBR is also affected by SOA. The E-V BCI system reaches the highest PBR in SOA for 125 ms.

Table 4 shows the effect of the SOA on the accuracy and PBR. Bold represents the highest value of the item. The $ACC_{visual\ speller}$ reaches the highest of 95.83% when the SOA is 175 ms, and the $ACC_{video\ watching}$ reaches the highest of 98.96% when the SOA is 125 ms. For the E-V BCI, both the accuracy and PBR reach the highest when the SOA is 125 ms, and they are 95.83% and 57.17 bit/min, respectively.

E. THE REPORTS OF VISUAL FATIGUE

After the subjects finished the experiment, we collected their visual fatigue report. The report records two feelings of subjects that what do you feel is the impact of SOA values on visual fatigue and what is the difference between the experimental feeling of asynchronous BCI and synchronous BCI. Most subjects felt that the smaller the SOA, the more fatigued they were. And they also felt more comfortable in asynchronous system than synchronous one because they were free to choose the time of output instruction.

IV. DISCUSSION

A. THE SOA AFFECTS BOTH THE ACCURACY AND SPEED OF THE E-V BCI SYSTEM

The SOA is positively related to the amplitudes of N200 and P300 but negatively related to the spectrum of the VEP.

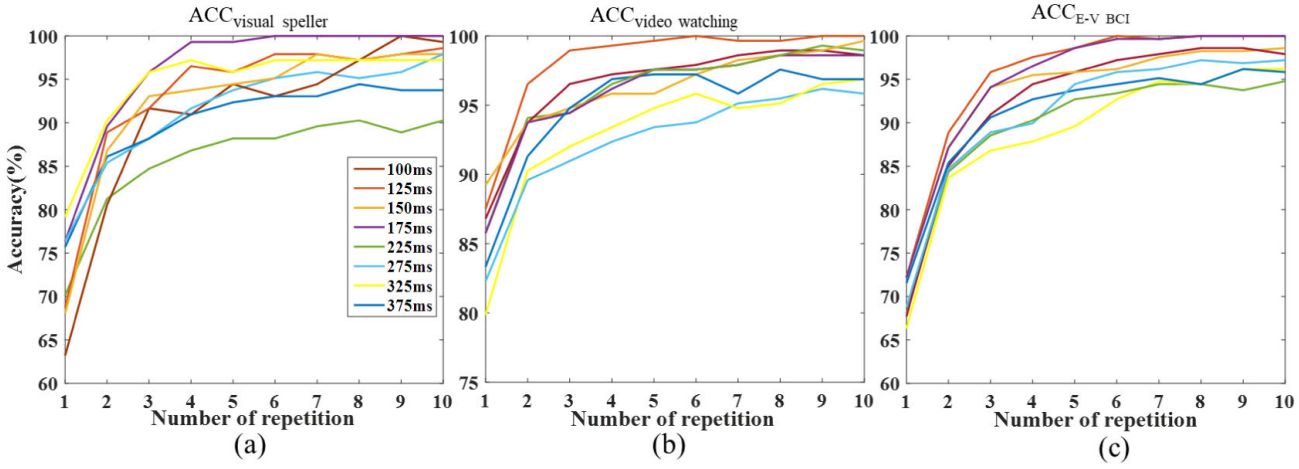

FIGURE 6. The accuracy of the BCI system.

TABLE 4. The accuracy and PBR.

	100ms	125ms	150ms	175ms	225ms	275ms	325ms	375ms
ACC $_{\text{visual speller}}$ (%)	91.67	91.67	93.06	**95.83**	84.72	88.19	95.83	88.19
ACC $_{\text{video watching}}$ (%)	96.53	**98.96**	94.79	94.44	94.44	90.97	92.01	94.79
$\text{ACC}_{\text{E-V BCI}}$ (%)	90.97	**95.83**	94.06	94.09	88.54	88.89	86.81	90.63
$\text{PBR}_{\text{E-V BCI}}$ (bits/min)	56.28	**57.17**	45.27	39.78	27.81	23.69	19.29	18.51

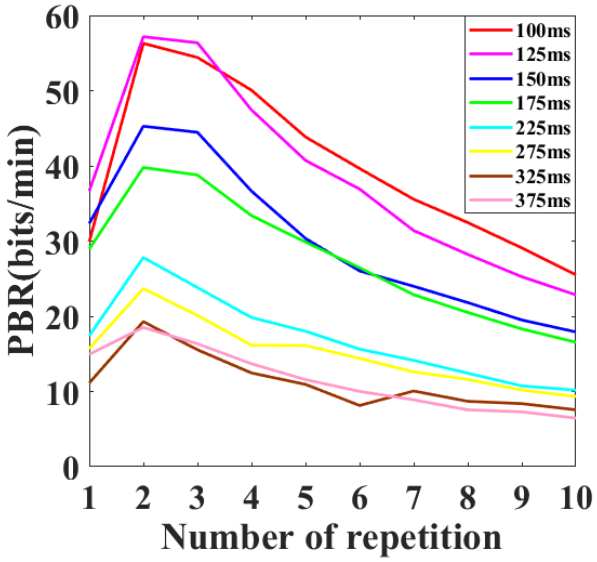

FIGURE 7. The PBR of the E-V BCI.

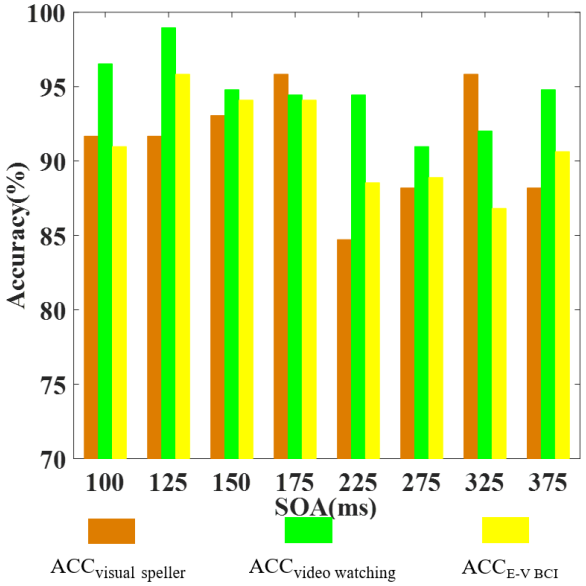

FIGURE 8. The accuracy under three repetition.

Therefore, when the SOA increases, the accuracy of identifying the target stimulus increases, while the accuracy of discriminating the brain states decreases. The integrated accuracy, the accuracy of discriminating the target stimulus and the two states, is more affected by the accuracy of discriminating the two states. Figure 6 shows that the tendency of $\text{Acc}_{\text{E-V BCI}}$ is similar to that of $\text{Acc}_{\text{video watching}}$ but not to that of $\text{Acc}_{\text{visual speller}}$. This is possibly because after the fusion of ERP and VEP, the classification result of VEP is more dominant than that of ERP in the recognition of brain state.

The SOA has a direct impact on the speed of outputting commands, which is represented by the PBR. Figure 7 shows that as the SOA increasing, the PBR is generally a decreasing trend. In addition, the SOA is the time interval between the stimuli. Therefore, increasing it leads to spending longer time completing a trial.

The difference between the ERP-based and VEP-based recognitions results in a non-linear relationship between the performances of the E-V BCI and the SOA. Figure 8 shows that with the increase in SOA, $\text{Acc}_{\text{visual speller}}$ is a trend that first increases, then decreases, and then increases. $\text{Acc}_{\text{video watching}}$ and $\text{Acc}_{\text{E-V BCI}}$ tend to increase first and then decrease. This is because the $\text{Acc}_{\text{E-V BCI}}$ is a combination of $\text{Acc}_{\text{visual speller}}$ and $\text{Acc}_{\text{video watching}}$ but is more similar to $\text{Acc}_{\text{video watching}}$. When the SOA is 125 ms, the $\text{Acc}_{\text{E-V BCI}}$ reaches the best value, which is 95.83%. $\text{Acc}_{\text{visual speller}}$ is 91.67% and $\text{Acc}_{\text{video watching}}$ is 98.96%. PBR is 57.17 bits/min. Therefore, when the SOA is 125 ms, the E-V BCI achieves the best system performance.

B. A COMPARE WITH OTHER ASYNCHRONOUS SYSTEMS

The state identification of subject's state plays an important role in improving the asynchronous system performance. Many researchers combine potentials to identify the working state of asynchronous systems. Santamaría-Vázquez *et al.* detected the SSVEP elicited by peripheral stimuli of ERP-based spellers to achieve asynchronous control, and reached an average accuracy of 95.5% for control state detection [34]. Zhang *et al.* proposed an individualized frequency band based optimized complex network (IFB-OCN) and to recognize the control and idle states for an asynchronous SSVEP system, with an average accuracy of 93.5 % [48]. Compared to them, we used the odd-ball paradigm to design the asynchronous BCI system, and used ERP and TSVEP to identify the working status and target. Our accuracy reached 95.83%, which is higher than them. Our asynchronous system has advantages in identifying system states and target, which provides a method for designing high-performance asynchronous BCI system.

V. CONCLUSION

This study sets eight stimulus intervals to study the effect of the SOAs on the asynchronous ERP and VEP-based BCI. Increasing the SOA can increase the amplitude of N200 and P300, but changes the dominant frequency of the VEP. The contradiction between identifying target and identifying the brain states makes the accuracy of asynchronous BCI system varies non-linear as SOA variation. Our study demonstrates that the SOA of 125 ms is optimal value for asynchronous BCI system to yield high accuracy and speed, which gives us insights for designing high-performance asynchronous ERP BCI.

In Fig 5 in frequency domain, besides the overall outlook, it would be more insightful that we extract specific frequency bands as features. The data of delta, theta and alpha frequency bands could be as feature vectors to distinguish target and non-target. In the future, we would use these frequency bands as the feature vector to improve the accuracy.

ACKNOWLEDGMENT

The authors sincerely thank all the participants for their voluntary participation and the reviewers as well as editors for their precious suggestions and comments. *(Mengfan Li and Minghong Gong contributed equally to this work.)*

REFERENCES

[1] J. Tang, M. Xu, Z. Liu, J. Qiao, S. Liu, S. Chen, T.-P. Jung, and D. Ming, "A brain-computer interface based on multifocal SSVEPs detected by inter-task-related component analysis," *IEEE Access*, vol. 8, pp. 138539–138550, 2020.

[2] F. Duan, D. Lin, W. Li, and Z. Zhang, "Design of a multimodal EEG-based hybrid BCI system with visual servo module," *IEEE Trans. Auto. Mental Develop.*, vol. 7, no. 4, pp. 332–341, Dec. 2015.

[3] Z. Zhang, F. Duan, J. Sole-Casals, J. Dinares-Ferran, A. Cichocki, Z. Yang, and Z. Sun, "A novel deep learning approach with data augmentation to classify motor imagery signals," *IEEE Access*, vol. 7, pp. 15945–15954, 2019.

[4] W. Li, F. Duan, and C. Xu, "Design and performance evaluation of a simple semi-physical human-vehicle collaborative driving simulation system," *IEEE Access*, vol. 7, pp. 31971–31983, 2019.

[5] S. L. Norman, D. J. McFarland, A. Miner, S. C. Cramer, E. T. Wolbrecht, J. R. Wolpaw, and D. J. Reinkensmeyer, "Controlling pre-movement sensorimotor rhythm can improve finger extension after stroke," *J. Neural Eng.*, vol. 15, no. 5, Oct. 2018, Art. no. 056026.

[6] X. Chen, B. Zhao, Y. Wang, and X. Gao, "Combination of high-frequency SSVEP-based BCI and computer vision for controlling a robotic arm," *J. Neural Eng.*, vol. 16, no. 2, Apr. 2019, Art. no. 026012.

[7] Á. Fernández-Rodríguez, F. Velasco-Álvarez, M. T. Medina-Juliá, and R. Ron-Angevin, "Evaluation of emotional and neutral pictures as flashing stimuli using a P300 brain–computer interface speller," *J. Neural Eng.*, vol. 16, no. 5, Sep. 2019, Art. no. 056024.

[8] J. Jin, Z. Chen, R. Xu, Y. Miao, X. Wang, and T.-P. Jung, "Developing a novel tactile P300 brain-computer interface with a cheeks-stim paradigm," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 9, pp. 2585–2593, Sep. 2020.

[9] H. Wu, H. He, and F. Tian, "Study of recognition algorithm based on transient visual evoked potential," *Transducer Microsyst. Technol.*, vol. 31, no. 4, pp. 47–49, 2012.

[10] B. Rebsamen, C. Guan, H. Zhang, C. Wang, C. Teo, M. H. Ang, and E. Burdet, "A brain controlled wheelchair to navigate in familiar environments," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 18, no. 6, pp. 590–598, Dec. 2010.

[11] Z. Feng, Q. He, J. Zhang, L. Wang, X. Zhu, and M. Qiu, "A hybrid BCI system based on motor imagery and transient visual evoked potential," *Multimedia Tools Appl.*, vol. 79, nos. 15–16, pp. 10327–10340, Apr. 2020.

[12] A. Kubacki and A. Jakubowski, "Controlling the industrial robot model with the hybrid BCI based on EOG and eye tracking," in *Proc. CMES*, 2018, Art. no. 020032.

[13] S. Sung, H. Huang, and C. Lin, "Development of EEG brain-computer interface system for control of shoulder-elbow rehabilitation robot," *J. Chin. Soc. Mech. Eng.*, vol. 41, no. 1, pp. 39–47, 2020.

[14] M. Li, W. Li, and H. Zhou, "Increasing N200 potentials via visual stimulus depicting humanoid robot behavior," *Int. J. Neural Syst.*, vol. 26, no. 1, Feb. 2016, Art. no. 1550039.

[15] M. Li, G. Yang, and G. Xu, "The effect of the graphic structures of humanoid robot on N200 and P300 potentials," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 9, pp. 1944–1954, Sep. 2020.

[16] W. Speier, C. Arnold, and N. Pouratian, "Integrating language models into classifiers for BCI communication: A review," *J. Neural Eng.*, vol. 13, no. 3, pp. 3–12, 2016.

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

[18] Y. Yu, Z. Zhou, Y. Liu, J. Jiang, E. Yin, N. Zhang, Z. Wang, Y. Liu, X. Wu, and D. Hu, "Self-paced operation of a wheelchair based on a hybrid brain-computer interface combining motor imagery and P300 potential," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 12, pp. 2516–2526, Dec. 2017.

[19] R. C. Panicker, S. Puthusserypady, and Y. Sun, "An asynchronous P300 BCI with SSVEP-based control state detection," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 6, pp. 1781–1788, Jun. 2011.

[20] M. Gong, M. Li, G. Xu, and F. Lin, "Research on identification method of brain working and idle state of asynchronous BCI," *J. Electron. Meas. Instrum.*, vol. 34, no. 4, pp. 11–19, 2020.

[21] M. Gong, G. Xu, M. Li, and F. Lin, "An idle state-detecting method based on transient visual evoked potentials for an asynchronous ERP-based BCI," *J. Neurosci. Methods*, vol. 337, May 2020, Art. no. 108670.

[22] M. Sugi, Y. Hagimoto, I. Nambu, A. Gonzalez, Y. Takei, S. Yano, H. Hokari, and Y. Wada, "Improving the performance of an auditory brain-computer interface using virtual sound sources by shortening stimulus onset asynchrony," *Frontiers Neurosci.*, vol. 12, p. 108, Feb. 2018.

[23] E. Donchin and M. G. H. Coles, "Is the P300 component a manifestation of context updating?" *Behav. Brain Sci.*, vol. 11, pp. 357–374, Sep. 1988.

[24] B. Z. Allison and J. A. Pineda, "Effects of SOA and flash pattern manipulations on ERPs, performance, and preference: Implications for a BCI system," *Int. J. Psychophysiol.*, vol. 59, no. 2, pp. 127–140, Feb. 2006.

[25] J. Jin, E. W. Sellers, and X. Wang, "Targeting an efficient target-to-target interval for P300 speller brain–computer interfaces," *Med. Biol. Eng. Comput.*, vol. 50, no. 3, pp. 289–296, Mar. 2012.

[26] H. Hong, L. Yan, and Z. Jin, "High-frequency SSVEP-based navigation of a humanoid robot," *Int. J. Innov. Comput. I*, vol. 45, no. 5, pp. 513–520, 2016.

[27] M. Daisuke, "A study of a SSVEP-based BCI using high frequency ficker stimulus," *IEICE Tech. Rep. Me Bio Cybern.*, vol. 112, no. 220, pp. 5–8, Sep. 2012.

[28] J. Hohne and M. Tangermann, "How stimulation speed affects event-related potentials and BCI performance," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2012, pp. 1802–1805.

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

[30] J. Jin, S. Li, I. Daly, Y. Miao, C. Liu, X. Wang, and A. Cichocki, "The study of generic model set for reducing calibration time in P300-based brain–computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 1, pp. 3–12, Jan. 2020.

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

[32] D. J. Krusienski, E. W. Sellers, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "Toward enhanced P300 speller performance," *J. Neurosci. Methods*, vol. 167, no. 1, pp. 15–21, Jan. 2008.

[33] J. Jin, E. W. Sellers, and Y. Zhang, "Whether generic model works for rapid ERP-based BCI calibration," *J. Neurosci. Methods*, vol. 212, no. 1, pp. 94–99, 2013.

[34] E. Santamaria-Vazquez, V. Martinez-Cagigal, J. Gomez-Pilar, and R. Hornero, "Asynchronous control of ERP-based BCI spellers using steady-state visual evoked potentials elicited by peripheral stimuli," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 9, pp. 1883–1892, Sep. 2019.

[35] Y. Xue, J. Tang, F. He, M. Xu, and H. Qi, "Improve P300 speller performance by changing stimulus onset asynchrony (SOA) without retraining the subject-independent model," *IEEE Access*, vol. 7, pp. 134137–134144, 2019.
[36] G. Xu, Y. Wu, and M. Li, "The study of influence of sound on visual ERP-based brain computer interface," *Sensors*, vol. 20, no. 4, p. 1203, Feb. 2020.
[37] H. Riechmann, A. Finke, and H. Ritter, "Using a cVEP-based brain-computer interface to control a virtual agent," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 6, pp. 692–699, Jun. 2016.
[38] B. Wu, Y. Su, and J. Zhang, "A virtual Chinese keyboard BCI system based on P300 potentials," *Acta Electronica Sinica*, vol. 37, no. 8, pp. 1732–1733, 2009.
[39] Z. Oralhan, "3D input convolutional neural networks for P300 signal detection," *IEEE Access*, vol. 8, pp. 19521–19529, 2020.
[40] Z. Oralhan, "A new paradigm for region-based P300 speller in brain computer interface," *IEEE Access*, vol. 7, pp. 106617–106626, 2019.
[41] A. Craik, Y. He, and J. L. Contreras-Vidal, "Deep learning for electroencephalogram (EEG) classification tasks: A review," *J. Neural Eng.*, vol. 16, no. 3, Jun. 2019, Art. no. 031001.
[42] A. Chamanzar, M. Shabany, A. Malekmohammadi, and S. Mohammadinejad, "Efficient hardware implementation of real-time low-power movement intention detector system using FFT and adaptive wavelet transform," *IEEE Trans. Biomed. Circuits Syst.*, vol. 11, no. 3, pp. 585–596, Jun. 2017.
[43] K. Chen, Q. Liu, and Q. S. Ai, "Multi-channel SSVEP pattern recognition based on MUSIC," *Appl. Mech. Mater.*, vol. 539, pp. 84–88, Jul. 2014.
[44] X. Chen, Y. Wang, S. Zhang, S. Xu, and X. Gao, "Effects of stimulation frequency and stimulation waveform on steady-state visual evoked potentials using a computer monitor," *J. Neural Eng.*, vol. 16, no. 6, Oct. 2019, Art. no. 066007.
[45] J. Jin, B. Z. Allison, T. Kaufmann, A. Kübler, Y. Zhang, X. Wang, and A. Cichocki, "The changing face of P300 BCIs: A comparison of stimulus changes in a P300 BCI involving faces, emotion, and movement," *PLoS ONE*, vol. 7, no. 11, Nov. 2012, Art. no. e49688.
[46] R. Cheng, R. W. Doerge, and J. Borevitz, "Novel resampling improves statistical power for multiple-trait QTL mapping," *G3-Genes Genomes Genet.*, vol. 7, no. 3, pp. 813–822, Mar. 2017.
[47] J. J. Green, J. A. Conder, and J. J. McDonald, "Lateralized frontal activity elicited by attention-directing visual and auditory cues," *Psychophysiology*, vol. 45, no. 4, pp. 579–587, Jul. 2008.
[48] W. Zhang, T. Zhou, J. Zhao, B. Ji, and Z. Wu, "Recognition of the idle state based on a novel IFB-OCN method for an asynchronous brain-computer interface," *J. Neurosci. Methods*, vol. 341, Jul. 2020, Art. no. 108776.

MENGFAN LI (Member, IEEE) received the B.Sc. degree from the Hebei University of Technology, in 2011, and the M.Sc. and Ph.D. degrees from Tianjin University, in 2017. She is currently an Assistant Professor with the Hebei University of Technology. Her main research interest includes brain–robot interaction.

GUANG YANG received the B.Sc. degree from Beihua University, in 2018. He is currently pursuing the M.Sc. degree with the Hebei University of Technology. His main research interest includes brain–computer interface.

ZHEN LIU is currently pursuing the B.Sc. degree with the Hebei University of Technology. His main research interest includes brain–machine interface.

MINGHONG GONG received the B.Sc. degree from the Shandong University of Science and Technology, in 2017, and the M.Sc. degree from the Hebei University of Technology, in 2020. His main research interest includes brain–robot interface.

GUIZHI XU (Member, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees from the Hebei University of Technology, in 1983, 1999, and 2002, respectively. She is currently a Professor with the Hebei University of Technology. Her main research interest includes biomedical electromagnetic technology.

FANG LIN received the B.Sc. degree from the Beijing University of Aeronautics and Astronautics, in 2015, and the M.Sc. degree from the Hebei University of Technology, in 2019. His main research interest includes brain–computer interface.