Enhancing Detection of Control State for High-Speed Asynchronous SSVEP-BCIs Using Frequency-Specific Framework

Yufeng Ke, *Member, IEEE*, Jiale Du, Shuang Liu, and Dong Ming, *Member, IEEE*

***Abstract*— This study proposed a novel frequency-specific (FS) algorithm framework for enhancing control state detection using short data length toward high-performance asynchronous steady-state visual evoked potential (SSVEP)-based brain-computer interfaces (BCI). The FS framework sequentially incorporated task-related component analysis (TRCA)-based SSVEP identification and a classifier bank containing multiple FS control state detection classifiers. For an input EEG epoch, the FS framework first identified its potential SSVEP frequency using the TRCA-based method and then recognized its control state using one of the classifiers trained on the features specifically related to the identified frequency. A frequency-unified (FU) framework that conducted control state detection using a unified classifier trained on features related to all candidate frequencies was proposed to compare with the FS framework. Offline evaluation using data lengths within 1 s found that the FS framework achieved excellent performance and significantly outperformed the FU framework. 14-target FS and FU asynchronous systems were separately constructed by incorporating a simple dynamic stopping strategy and validated using a cue-guided selection task in an online experiment. Using averaged data length of 591.63±5.65 ms, the online FS system significantly outperformed the FU system and achieved an information transfer rate, true positive rate, false positive rate, and balanced accuracy of 124.95±12.35 bits/min, 93.16±4.4%, 5.21±5.85%, and 92.89±4.02%, respectively. The FS system was also of higher reliability by accepting more correctly identified SSVEP trials and rejecting more wrongly identified ones. These results suggest that the FS framework has great potential to enhance the control state detection for high-speed asynchronous SSVEP-BCIs.**

***Index Terms*— Asynchronous brain-computer interface, control state detection, steady-state visually evoked potentials (SSVEP).**

I. Introduction

BRAIN-COMPUTER interface (BCI) can establish direct links between the human brain and external devices like computers by decoding the user's intention from neural signals [1], [2], [3], [4]. BCIs have shown great potential to provide muscle-independent communication channels and to help disabled and abled people to restore or enhance their abilities. According to users' autonomy during using BCI systems, a BCI can be classified as synchronous or asynchronous. An asynchronous BCI allows users to communicate freely by using it whenever they want. In contrast, a synchronous BCI only enables users to operate at a particular predefined time and usually needs a synchronization cue for the beginning of each trial. As one of the most popular noninvasive BCI paradigms, steady-state visually evoked potential (SSVEP)-based BCIs have been heatedly researched in the past decades and have demonstrated the advantages of less training and higher information transfer rate (ITR) using efficient feature extraction or machine learning methods like the extended canonical correlation analysis (CCA) [5], task-related component analysis (TRCA) [6], correlated component analysis [7], task-discriminant component analysis [8], and deep learning-based [9], [10], [11], [12] target identification algorithms.

However, limitations on users' autonomy during the interaction are still one of the challenges for user-friendly and practical applications in the existing high-performance SSVEP-BCI studies because of their synchronous manner. In contrast to the high performance (high ITR, high accuracy, and a large number of targets) and the large number of studies on synchronous SSVEP-BCI, the asynchronous SSVEP-BCI has been understudied and exhibited unsatisfactory performance. The critical point of establishing an asynchronous SSVEP-BCI is distinguishing users' 'control state.' A user may be in either the intentional control (IC) state or the non-control (NC) (or idle) state, according to the user's gazing state (gazing at an SSVEP flicker or not). The key challenge of an

Manuscript received 10 September 2022; revised 17 January 2023; accepted 13 February 2023. Date of publication 16 February 2023; date of current version 24 February 2023. This work was supported in part by the National Key Research and Development Program of China under Grant 2021YFF1200603 and in part by the National Natural Science Foundation of China under Grant 61806141. *(Yufeng Ke and Jiale Du contributed equally to this work.) (Corresponding authors: Shuang Liu; Dong Ming.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee of Tianjin University, Tianjin, China (under Application No. TJUE-2022-047).

Yufeng Ke, Shuang Liu, and Dong Ming are with the Academy of Medical Engineering and Translational Medicine, Tianjin University, Tianjin 300072, China (e-mail: clarenceke@tju.edu.cn; shuangliu@tju.edu.cn; richardming@tju.edu.cn).

Jiale Du is with the College of Precision Instruments and Optoelectronics Engineering, Tianjin University, Tianjin 300072, China (e-mail: jaledu@tju.edu.cn).

This article has supplementary downloadable material available at https://doi.org/10.1109/TNSRE.2023.3246359, provided by the authors.

Digital Object Identifier 10.1109/TNSRE.2023.3246359

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

asynchronous SSVEP-BCI is to distinguish between the IC and NC states effectively. Based on the method used to detect the change in the user's control state, there have been two main types of asynchronous SSVEP-BCI research: i) establishing a hybrid BCI (hBCI) by combining an SSVEP-BCI and switches based on other physiological signals and ii) directly classifying between IC and NC states using SSVEP features.

Several hBCI studies that employed electrooculogram (EOG) [13], [14], [15], [16], electromyogram (EMG) [16], [17], P300 [18], [19], and respiration signals [20] to establish switches for SSVEP-BCIs have shown their potential to construct asynchronous SSVEP-BCIs for practical applications. In these studies, the users are usually required to perform intentional actions like eye blinks, occlusal movements, P300 tasks, and breath-holding to produce expected patterns to trigger the switch. Although users can turn the SSVEP-BCI on/off at any time in this kind of system, additional mental resources or muscle abilities are required, thus limiting its ease of use and application scenarios, especially for disabled patients.

In addition to the hBCI solutions, researchers have attempted to distinguish between the IC and NC states using SSVEP features. Xia et al. designed a four-target asynchronous SSVEP-BCI by combining the thresholding ratio of canonical correlation analysis (CCA) coefficients with a sliding window strategy [21]. Some other studies have also employed similar methods to distinguish between the IC and NC states, although there were some differences in the implementation approach of the thresholding methods [22], [23], [24]. In a recent study, Chen et al. combined a dynamic window method and a thresholding strategy for asynchronously controlling a robotic arm in an augmented reality (AR) environment [25]. In an online robot movement task, Chen et al. achieved an average SSVEP recognition accuracy of 92.49% at the average time cost of 2.04 s per command selection. Although the thresholding method can provide a training-free approach to recognizing the control state of users, the recognition accuracy and the response speed of these methods need to be improved, especially in a system of many candidate targets. Another recent study proposed a training-free spatio-temporal equalization multi-window algorithm based on the 'statistical inspection-rejection decision' mode to improve performance for a forty-target system [26]. In an online asynchronous spelling task, the system achieved an average accuracy of 97.2% but costed a relatively long average input time (3.06 s) for a single character, although a dynamic-window approach was employed.

Other researchers have attempted to classify IC and NC states using machine learning methods. Poryzala et al. implemented an IC detection method by clustering analysis of CCA coefficients after CCA-based spatial filtering of multichannel SSVEP responses, resulting in improved detection speed and accuracy [27]. In a three-target SSVEP-BCI study by Zhang et al., FFT features of SSVEP responses at the frequencies of both the attended and the unattended stimuli were employed to construct three binary classifiers based on support vector machine (SVM) for recognition of the three frequencies and the idle state [28]. Subsequently, they identified the three targets and the idle state by comparing the maximum output of the three classifiers to a predefined threshold and achieved an average control state detection accuracy of 88.0% and average idle state false positive rates (FPRs) ranging from 7.4% to 14.2% using 3s EEG data. Using a method similar to [28] in identifying the control state, Han et al. implemented a ten-target asynchronous system and achieved an averaged frequency recognition accuracy of 90% and an average FPR of 15% using 2 s EEG data by improving the feature extraction method [29]. Different from the previous two studies, which constructed binary classifiers to distinguish between each frequency and the idle state, Suefusa et al. trained a multi-class SVM classifier with correlation coefficient features extracted by the multiset-CCA algorithm to classify between all the targets and the idle state in a 28-target mixed frequency and phase-coded SSVEP system [30]. As the criterion of the performance regarding discriminability of the IC/NC states in this study, the area under the receiver operating characteristic curve (AUC-ROC) reached 0.919 using the data length of 3 s. With the rapid development of deep learning technology, typical deep learning methods like convolutional neural networks (CNN) have also been introduced into asynchronous SSVEP-BCI research. An FFT-CNN-CCA pipeline that firstly distinguished between IC and NC by a CNN using FFT features and subsequently recognized the target for the IC states using the CCA-based method achieved an average ACC of 90.09% in IC/NC classification using a data length of 2 s in a four-target system [31]. Ravi et al. classified NC and four IC states directly using a CNN architecture with complex FFT features and achieved an average accuracy of 82% using 1s-long EEG segments [32].

The studies mentioned above in the last decade have achieved impressive IC/NC classification performance in accuracy. However, whether training-free or machine learning-based, these studies occasionally required data lengths of almost 2 s or longer to achieve a high accuracy rate above 90%. It has already been a consensus in the SSVEP-BCI community that the overall performance of a system involves a trade-off between accuracy and speed. Generally, the longer the length of SSVEP data, the higher the recognition accuracy and the slower the response speed. Thus, it is still a challenge for the methods in the existing studies to simultaneously provide sufficient recognition accuracy and a fast response speed for time-critical application scenarios. There is still a requisite for a faster and more accurate algorithm framework for control state detection toward high-performance asynchronous SSVEP-BCI systems.

This study is dedicated to proposing an efficient frequency-specific (FS) algorithm framework to classify control states (IC/NC) for asynchronous SSVEP-BCI using FS features and classifiers. The FS algorithm framework first identified the potential target frequency using the TRCA-based template matching method and subsequently distinguished between IC (the identified frequency) and NC states using classifiers trained on the features specifically related to the identified target frequency. A frequency-unified (FU) algorithm

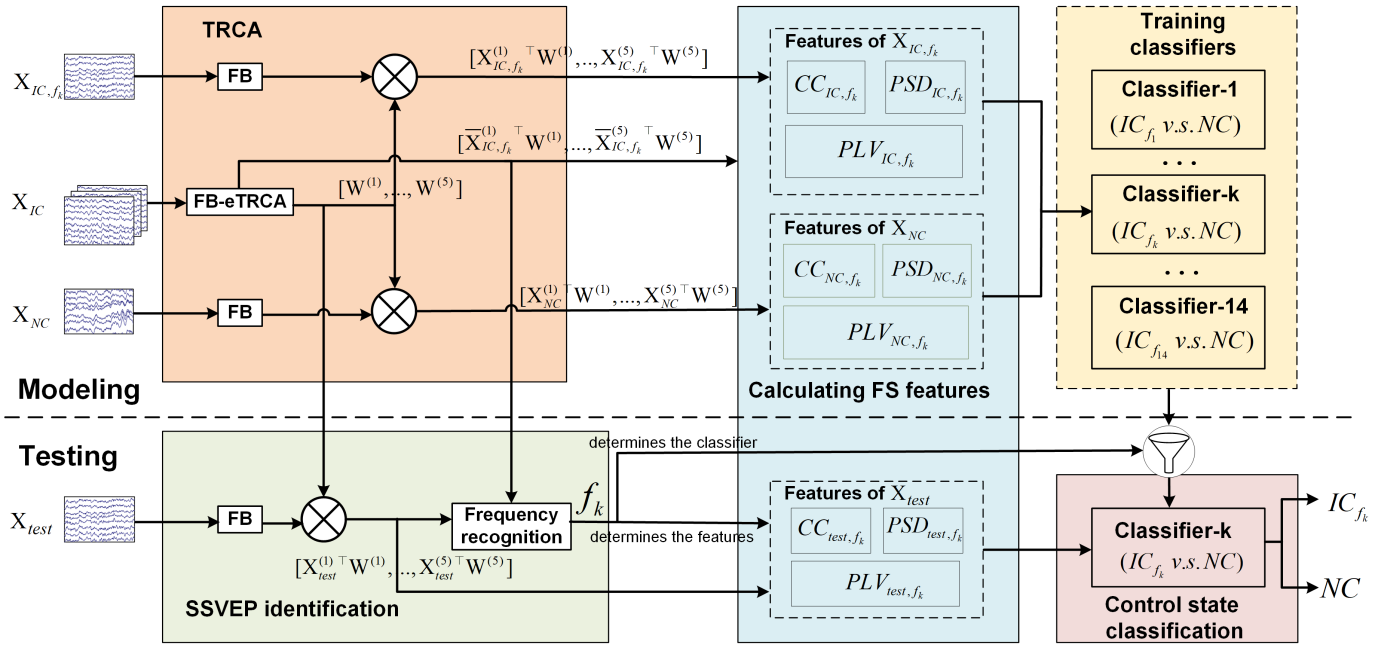

Fig. 1. Diagram of the proposed FS framework. X_{IC} and X_{NC} denote the training data from the IC and NC states, respectively. X_{IC,f_k} denotes the SSVEP data of the k^{th} stimulus frequency f_k (k=1,..., 14). X_{test} denotes the test data. $W^{(m)}$ ($m = 1, \ldots, 5$) denotes the spatial filters, and $\bar{X}_{IC}^{(m)}$ denotes the corresponding SSVEP templates, of m^{th} sub-band obtained through eTRCA from X_{IC} after FB analysis. $X_{IC,f_k}^{(m)}$, $X_{NC}^{(m)}$ and $X_{test}^{(m)}$ denote the m^{th} sub-band obtained from X_{IC,f_k}, X_{NC} and X_{test}, through FB analysis, respectively.

framework that distinguished between IC and NC states using a classifier trained on SSVEP features related to all candidate frequencies was proposed to compare with the FS algorithm framework. Different feature extraction methods were examined within the two algorithm frameworks in pseudo-online analyses using short data lengths. Several performance metrics were compared between the two algorithm frameworks and among three feature extraction approaches. Finally, online FS and FU-based asynchronous systems were implemented separately by incorporating the FS and FU frameworks with a simple dynamic stopping strategy. The performance of the online systems was validated using a cue-guided selection task in an online experiment.

II. Methods

For asynchronous SSVEP-BCIs, the system's overall performance depends on both SSVEP identification and control state classification. In this study, both the FS and FU algorithm frameworks used the same SSVEP identification method but differed in their approach to classifying control states. This section presents details of the frequency identification algorithm and the two algorithm frameworks.

A. SSVEP Identification

The ensemble TRCA (eTRCA)-based method [6], [33] was employed for SSVEP identification in both the FS and FU frameworks. The filter bank (FB) analysis [34] was also applied to EEG epochs by using zero-phase Chebyshev Type I infinite impulse response (IIR) filters. EEG epochs were decomposed into 5 sub-bands by filters with pass-bands of [(m∗8-2) Hz, 80 Hz], where $m = 1, 2, \ldots, 5$. The template matching-based method was employed to calculate the correlation coefficient (CC) features between SSVEP templates and test epochs for each sub-band. The potential target frequency for a test epoch was determined by the weighted sum of squares of the CC features of the 5 sub-bands, just as has been done in [6].

B. Control State Detection Algorithm Frameworks

Fig. 1 shows the proposed FS control state detection algorithm framework. As for comparison, an FU algorithm framework, as shown in Fig. 2, has also been implemented. It is definite that both algorithm frameworks identified SSVEP and used the eTRCA in the same way, but they were different in calculating features, training classifier(s), and applying the classifier(s) for control state detection.

In the SSVEP identification procedure, the SSVEP data, X_{IC}, recorded in IC state were used to obtain the TRCA spatial filters and the SSVEP templates. In the control state detection procedure, features extracted from the data of IC state, X_{IC}, and NC state, X_{NC}, were adopted to train the IC/NC classifier(s). Performance of SSVEP identification and IC/NC classification were evaluated on the test data, X_{test}. More detailed procedures for the two frameworks, feature extraction methods, and classifiers are as follows.

1) FS Framework: The FS framework sequentially incorporated the TRCA-based SSVEP identification and a classifier bank for control state detection. At the training stage, TRCA spatial filters, $[W^{(1)}, \ldots, W^{(5)}]$, and the SSVEP templates, $[\bar{X}_{IC}^{(1)^T} W^{(1)}, \ldots, \bar{X}_{IC}^{(5)^T} W^{(5)}]$, of all sub-bands were firstly obtained from the IC data, X_{IC}, through FB-eTRCA.

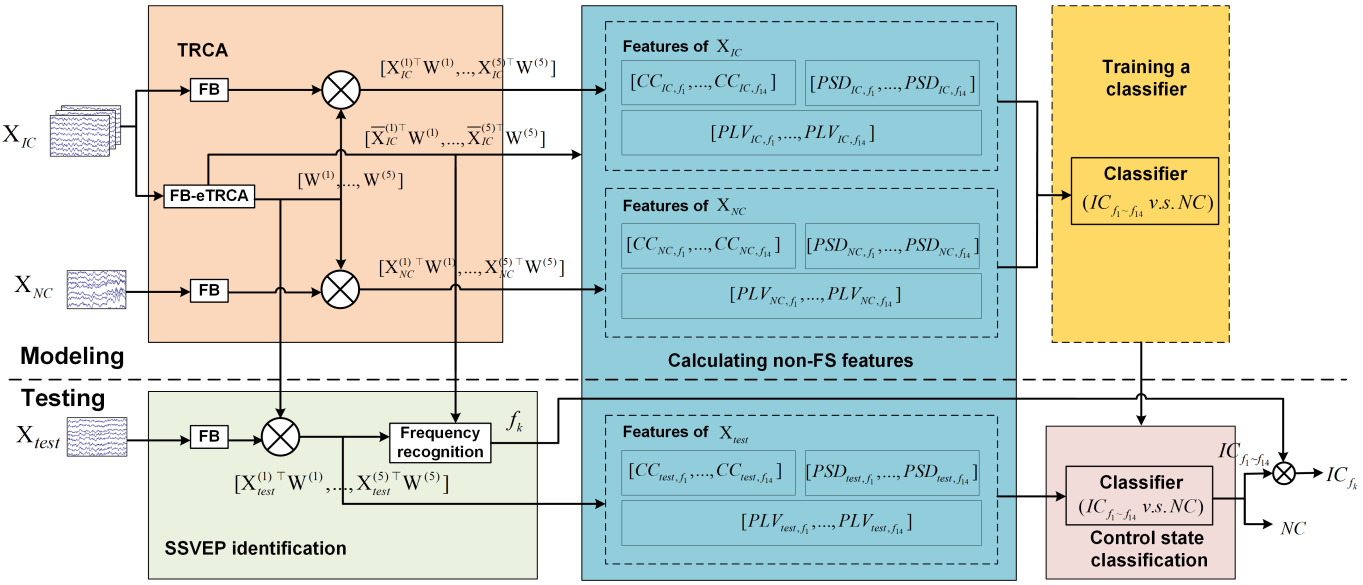

Fig. 2. Diagram of the FU framework.

Afterwards, the classifier bank containing fourteen (the total number of SSVEP targets) binary classifiers was constructed, each of the classifiers was trained on features only related to a specific candidate frequency calculated from spatially filtered data of the NC and IC states and thus could distinguish between the NC state and one of the fourteen SSVEPs. That is to say, the $k^{th}(k = 1, \ldots, K$, where K=14 is the number of targets) classifier was trained on k^{th} frequency (f_k)-related features calculated from X_{NC} and X_{IC,f_k}, the SSVEP data of f_k from IC state.

At the testing stage, the potential SSVEP frequency f_k of an EEG epoch, X_{test}, was firstly identified by applying the FB-eTRCA-based template matching approach, just as has been done by Nakanishi et al. [6]. Afterward, the f_k-related features were calculated from the spatially filtered data of each sub-band of X_{test} and then fed into the k-th classifier to determine whether the SSVEP frequency identified above was genuinely derived from SSVEP or not. Namely, the SSVEP identification result determined how the features would be calculated and which classifier would be applied to distinguish between IC and NC states.

2) FU Framework: The TRCA-based analysis and the SSVEP identification procedures of the FU framework were identical to those of the FS framework. The main difference was that only one unified binary classifier was constructed by training on features related to all of the candidate SSVEP frequencies calculated from data of X_{NC} and X_{IC} in the FU framework. At the testing stage, the IC or NC state of the testing data was determined by the only unified classifier independently of the SSVEP identification results. That is to say, SSVEP identification and control state detection were sequentially conducted in the FS framework, but they could be done simultaneously in the FU framework.

3) Feature Extraction Methods for Control State Detection: Three kinds of feature extraction methods, Pearson's correlation coefficient (CC), Fast Fourier Transform (FFT)-based Power Spectral Density (PSD), and Phase Locking Value (PLV), were adopted and compared between each other in this study. The CC features were extracted by calculating CCs between the spatially filtered SSVEP templates and the data for feature extraction for each sub-band after FB analysis. PSD features were extracted for the spatially filtered training and testing data at fundamental and harmonic ($h = 2 \sim 5$) frequencies by calculating the ratios between PSDs of these frequencies and the total power using the first sub-band. PLV features were extracted for training and testing data at fundamental and harmonic ($h = 2 \sim 5$) frequencies after spatial filtering by calculating PLVs between them and the SSVEP templates and also the corresponding cosine templates using the Hilbert transform-based phase extraction method [35]. As has been mentioned above, the difference between the feature calculation methods of the two frameworks was that the FS framework only extracted the features related to a specific SSVEP frequency for training one of the FS classifiers, while the FU framework extracted features of all the candidate SSVEP frequencies for training only one FU classifier. The detailed feature extraction process can be found in the supplementary materials.

4) Classifiers: In both frameworks, binary classifiers were adopted to discriminate between IC and NC states based on a classical low-cost machine learning method, the linear discriminant analysis (LDA) [36], since it is less demanding on sample size [37].

III. Experiments

In this study, offline and online experiments were conducted separately. The purpose of the offline experiment was to collect data for validating the performance of the control state detection algorithm frameworks and feature extraction methods using different data lengths. The online experiment validated the feasibility of a high-speed asynchronous SSVEP-BCI system using the proposed FS framework and compared the

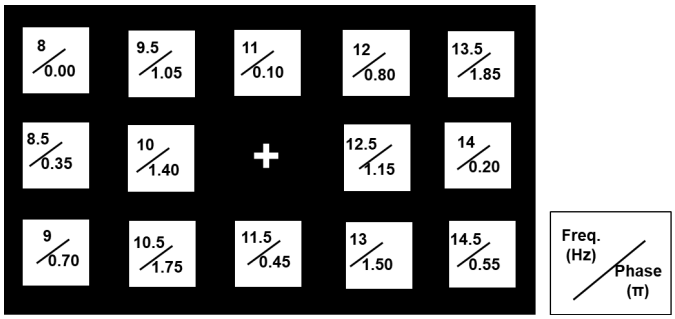

Fig. 3. Layout of the SSVEP stimuli.

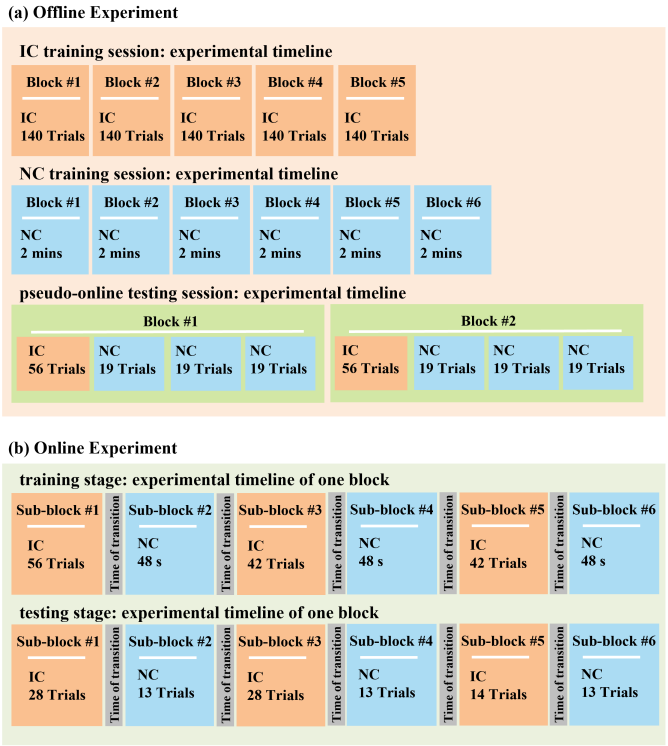

Fig. 4. Timeline schematic diagrams for the offline and online experiments.

performance between FS and FU frameworks under a cue-guided selection task. This section introduces the experimental environment, the detailed procedure of the two experiments, and the performance evaluation methods.

A. Experimental Environment

1) Subjects: Fifteen subjects (22-26 years old, incl. 7 males) and fourteen subjects (23-27 years old, incl. 5 males) participated in the offline and online experiments, respectively, with informed consent and payments. They all had a normal or corrected-to-normal vision and had no history of mental diseases or neurological disorders. The experimental procedures have been approved by the ethics committee of Tianjin University.

2) Stimulus Design: Fig. 3 illustrates the layout of the SSVEP stimuli displayed on a 27-inch Liquid Crystal Display (LCD) monitor with a refresh rate of 120 Hz. There were 14 flickering squares (170 × 170 pixels) coded using a joint frequency and phase modulation (JFPM) method [38] and a non-flickering cross in the center of the monitor. The flicker frequencies were set to 8 Hz ∼ 14.5 Hz with an interval of 0.5 Hz, and the initial phase interval was 0.35 π between adjacent frequencies. The stimulus programs of offline and online experiments were implemented using the Psychophysics Toolbox Version 3 [39].

3) EEG Recording Preprocessing: EEG data were recorded using the Compumedics Neuroscan Synamps2 hardware and Scan 4.5 software (Neuroscan, Inc.) with a sampling rate of 1000 Hz. Data from nine posterior channels (Pz, PO5, PO3, POz, PO4, PO6, O1, Oz, and O2, according to the extended 10-20 system) referenced to the vertex and grounded in the prefrontal region were used in the following analyses. Event triggers generated by the stimulus program at the onset of flickering were synchronized to the EEG data. Data epochs of IC and NC states were extracted according to the event triggers. All data epochs were band-pass filtered from 7 Hz to 90 Hz to preserve the first five harmonic components for all stimulation frequencies and notch filtered at 50 Hz to remove powerline noise by zero-phase forward and reverse filtering with Butterworth filters.

B. Offline Experiment

The offline experiment contained three sessions (as shown in Fig. 4a), an IC training session for SSVEP data collection, an NC training session for NC data collection, and a pseudo-online testing session including SSVEP and NC trials for testing data collection. To relieve subjects' visual fatigue, they were asked to have breaks for several minutes between consecutive experimental blocks. The detailed experimental process is described as follows:

1) The IC Training Session: This session contained five blocks, and each block included 140 trials. In each block, each of the squares was visually and randomly cued ten times in a trial-by-trial manner. At the beginning of each trial, subjects were instructed to turn their gaze to the cued square during the 0.5 s cue presentation interval. Then, the cued square started to flicker for 1 s, and subjects were instructed to gaze at it till the appearance of the next cue. After the flicker ended, there was a 1 s break before the cue of the subsequent trial. Therefore, this session lasted about 30 minutes, and each subject completed 700 trials with 50 trials for each frequency.

2) The NC Training Session: This session contained six 2-minute NC blocks and was conducted to simulate the NC state. In each block, subjects were randomly instructed to either gaze at the non-flickering white cross in the center of the monitor (on-screen NC), to look outside of the monitor and play with their cellphone (off-screen NC), or to close their eyes (eye-closed NC). Two blocks were conducted for each of the NC states.

3) The Pseudo-Online Testing Session: Subjects completed two blocks in the pseudo-online testing session. In each block, subjects completed a 56-trial IC sub-block followed by three 19-trial NC sub-blocks. The NC sub-blocks were randomly assigned as one of the three NC states. Auditory cues were employed at the beginning of each trial in this session,

considering the off-screen and eye-closed NC conditions. The timing of each trial was identical to that of the IC session.

For the offline experiment, EEG epochs in [0.14 s to $0.14+l$ s] were extracted from the IC session and the pseudo-online session according to the event triggers at the onset of flickering, where the time 0 s indicated stimulus onset and l was the data length and ranged in [0.5 1] with an interval of 0.1. The EEG data from the NC session were split into l s epochs without overlap. Subsequently, performance was validated on data lengths from 0.5 s to 1 s by training the algorithms on data from IC and NC training sessions and testing on data from the pseudo-online session. The number of samples for classifier training was balanced between IC and NC classes. Each of the fourteen classifiers in the FS framework was trained on fifty IC samples. As for the FU framework, the classifier was trained on 700 IC samples (14 frequencies $\times$ 50 samples/frequency).

C. Online Experiment

To validate the online performance of the two frameworks, asynchronous BCI systems based on the FS and FU frameworks were separately constructed. The dynamic stopping strategy was employed here for high-speed asynchronous BCI systems because previous studies have shown its potential to improve the reliability of SSVEP-BCI with short data lengths [33], [40]. The online experiment was divided into a training stage and a testing stage. EEG data were collected to train the TRCA-based and LDA-based algorithms for FS and FU systems in the training stage. The online performance of the two systems was validated separately via a cue-guided selection task in the testing stage.

1) Training Stage: The training stage consisted of 5 blocks, each including 3 IC sub-blocks and 3 NC sub-blocks (as shown in Fig. 4b). The sub-blocks were presented in the order of IC-NC-IC-NC-IC-NC. There were 4-s intervals between consecutive sub-blocks to allow subjects to switch between different states. In each block, the number of trials for the three IC sub-blocks were 56, 42, and 42, resulting in 140 IC trials and 10 trials for each target frequency. Thus, 50 trials for each target frequency were generated from the training stage. Each IC trial lasted 2 s, including 1 s for visual stimulation and 1 s for gaze shifting. As mentioned in the offline experiment, there were three different NC states: on-screen, off-screen, and eye-closed. Each of the 3 NC sub-blocks lasted 48 s and was randomly assigned as one of the three NC states. All of the NC sub-blocks were segmented into 1 s non-overlapping epochs. To avoid visual fatigue, subjects were asked to have breaks for several minutes between consecutive blocks.

The data acquired in the training stage were used as individual training data in the testing stage. To implement dynamic stopping asynchronous BCI systems, models of different data lengths of both the FS and the FU frameworks were trained for data lengths from 0.46 s to 0.98 s with an interval of 0.04 s because the shortest data length could be accessed online from the hardware was 0.04 s. This allowed the online systems to select a model trained on EEG epochs of $[0.46+n\times0.04]$ s ($n = 0, 1, \ldots, 13$; n representing the number of 0.04-s epochs received from the hardware) according to the length of the data obtained in real-time. Considering a latency delay in the visual pathway, the IC data epochs were extracted in [0.14 s $(0.6+n\times0.04)$ s] according to the onset of visual flickering. The NC epochs corresponding to the length of the IC epochs were directly extracted from the aforementioned 1-s epochs from the NC sub-blocks without considering the visual latency. The two frameworks' TRCA-based spatial filters and the SSVEP templates were constructed using all IC trials recorded in the training stage. The LDA-based FS classifiers and FU classifiers were trained on a balanced number of IC and NC samples and also a balanced number of the three NC samples.

2) Testing Stage: The testing stage of the cue-guided selection task included 4 blocks, 2 blocks were conducted to validate the FS framework, and the other 2 blocks were performed to validate the FU framework. The two conditions were randomly conducted, and subjects were not informed which algorithm frameworks were used in all 4 blocks. Each block consisted of 3 IC sub-blocks and 3 NC sub-blocks in the order of IC-NC-IC-NC-IC-NC (as shown in Fig. 4b). In each block, the number of trials for the three IC sub-blocks was 28, 28, and 14 in order, resulting in 70 IC trials in total and 5 trials for each target. Each target stimulus was randomly cued with a red cross presenting in the center of the stimulus for 0.8 s. The 3 NC sub-blocks were also randomly assigned as one of the three NC states, and each was divided into 13 NC trials. Thus, there were 140 IC and 78 NC trials for the FS and FU systems in the testing stage. When switching between different sub-blocks, voice prompts indicated whether the next sub-block would be an IC task (cue-guided selection) or one of the three NC states. There were 4-s intervals between consecutive sub-blocks to allow subjects to switch between different states.

To obtain more reliable output using short data lengths, this study implemented the dynamic stopping BCI systems by determining whether the identification results of two consecutive overlapping EEG epochs were the same. If the results of epochs [0.14 s to $0.46+n\times0.04$ s] and [0.14 s to $(0.46+(n+1)\times0.04)$ s] were identical, the systems would output the result and send a stopping signal to the visual stimulus program. If not, the systems would receive more data from the hardware and determine if the results of epochs [0.14 s to $(0.46+(n+1)\times0.04)$ s] and [0.14 s to $(0.46+(n+2)\times0.04)$ s] were identical, and so on. When the visual flickering of a trial reached 0.98 s and ended, if there were still not two consecutive overlapping epochs of identical identification result, the result of epoch [0.14 s to 0.98 s] would be the output of the trial.

Because the online experiment used a dynamic stopping strategy, the duration of each trial consisted of a fixed gaze shifting time of 0.8 s, a dynamic visual stimulation time of up to 0.98 s, and a fixed feedback presentation time of 1 s. Visual and auditory feedback was provided to the subjects in real-time. Once a target was selected and the stimulus program received feedback from the online processing program, the SSVEP stimulation stopped flickering. At the same time, a voice was prompted for the selected target and a hollow box was displayed around the selected stimulus. The cue for the next target appeared after the 1-s feedback duration. Subjects

TABLE I
A CONFUSION MATRIX OF EXPLANATION ABOUT THE FOUR DIFFERENT DECISIONS OF THE IC TRIALS

		Decision of control state detection	
		IC (acceptance)	NC (rejection)
Decision of frequency identification	Correct	Correct acceptance (CA)	Wrong rejection (WR)
	Wrong	Wrong acceptance (WA)	Correct rejection (CR)

were asked to rest for several minutes between consecutive blocks to avoid visual fatigue.

D. Performance Evaluation

1) Performance Evaluation for Offline Experiment: The area under the receiver operating characteristic curve (AUC-ROC), true positive rate (TPR), false positive rate (FPR), 15-class (fourteen SSVEP frequencies and the NC state) balanced accuracy, and confusion matrix were employed to quantify the performance of both frameworks. The present study regarded the IC states as positive and the NC as negative.

We further analyzed the final recognition results of the IC states in detail by considering both frequency identification and control state detection. As shown in Table I, there were four different possible outcomes for the IC trials, according to the frequency identification and control state detection decisions. If an SSVEP trial was classified as IC, the system would accept the output whether it was correctly identified or not. Otherwise, the result would be rejected. The accepted correctly identified trials were defined as the correct acceptances (CA), and the accepted wrongly identified trials were the wrong acceptances (WA). The rejected correctly identified trials were defined as the wrong rejections (WR), and the rejected wrongly identified trials were the correct rejections (CR). Then, the CA rate (CAR) and CR rate (CRR) were calculated as CA/(CA+WA) and CR/(WR+CR), respectively. For practical applications, classifying a misidentified SSVEP epoch into IC usually has more negative effects than classifying it as an NC class. So, we expected the FS framework to achieve a higher CAR and CRR than the FU framework.

The offline performance metrics were compared between the two frameworks and among three features at different data lengths from 0.5 s to 1 s with an interval of 0.1 s. Because the training and testing data were independent in this study, repeated-measure analysis of variance (rm-ANOVA) and paired t-test were employed at the group level to examine the effects of algorithm framework and feature under the premise of satisfying normality and variance homogeneity. When the premise was unsatisfied, the non-parametric bootstrap approach was employed with 1,000 repeated times of random resampling. The p-values corrected using the Bonferroni-Holm-based method [41] (p_{BH}) were reported for multiple comparisons.

2) Performance Evaluation for Online Experiment: SSVEP recognition accuracies, ITRs, TPRs, FPRs, average data lengths, CARs, CRRs, confusion matrixes, and balanced accuracies were reported and compared between the online FS and FU systems. The ITR [42] is calculated as:

$$\mathrm{ITR} = \left[\log_2 K + P\log_2 P + \log_2\left(\frac{1-P}{K-1}\right)\right] \times \frac{60}{T} \quad (1)$$

where $K = 14$ is the total number of targets, P is the SSVEP recognition accuracy, and T is the averaged time per selection consisting of the gaze shifting time, the visual latency, and the averaged data length used online. Paired t-tests were performed to compare these metrics between the two algorithm frameworks.

IV. RESULTS OF OFFLINE EXPERIMENT

A. Control State Detection Performance

Fig. 5 shows the averaged performance of the control state detection results across all subjects with different data lengths (l). The comparisons across algorithm frameworks and features indicate that the proposed FS framework outperformed the FU framework. The FS framework achieved the highest performance regardless of feature and data length. Two-way rm-ANOVAs were performed to analyse the effects of algorithm framework and feature on AUC-ROC, TPR, and FPR, separately for all data lengths.

1) AUC-ROC: There were significant interactions between the effects of the algrithm framework and feature at data lengths of 0.5 - 1 s ($Fs(2, 28) > 11.98$, $ps_{BH} < 0.001$). Simple main effects analyses showed that both algorithm framework ($Fs(1, 14) > 89.37$, $ps_{BH} <$ 10e-6) and feature ($Fs(2, 28) > 9.33$, $ps_{BH} < 0.001$) did have statistically significant effects on AUC-ROC at all data lengths. One-way rm-ANOVAs revealed significant effects of feature for AUC-ROC of FS at short data lengths (l = 0.5 s - 0.6 s: $Fs(2, 28) > 5.61$, $ps_{BH} < 0.05$) and of FU at data lengths of 0.5 - 1 s ($Fs(2, 28) > 11.00$, $ps_{BH} < 0.001$).

Furthermore, post hoc analyses using paired t-tests showed a statistical difference between the FS and FU frameworks for all features at all data lengths ($ts(14) > 4.53$, $ps_{BH} < 0.001$). The highest AUC-ROCs were achieved with CC features for the FS framework but with PSD features for the FU framework. At data lengths of 0.5 s and 1 s, the highest AUC-ROCs were 0.964±0.025 and 0.980±0.017 for the FS framework and 0.877±0.042 and 0.911±0.048 for the FU framework, respectively.

2) TPR and FPR: Two-way rm-ANOVAs revealed that there were statistically significant interactions between the effects of algorithm framework and feature for all data lengths (TPR: $Fs(2, 28) > 9.60$, $ps_{BH} < 0.001$; FPR: $Fs(2, 28) > 5.19$, $ps_{BH} < 0.05$). Main effects across algorithm frameworks on both TPR and FPR were revealed at all of the data lengths (TPR: $Fs(1, 14) > 32.74$, $ps_{BH} <$ 10e-5; FPR: $Fs(1, 14) > 127.71$, $ps_{BH} = 0.000$). Main effects across features were found for TPR at short data lengths (l = 0.5 - 0.8 s: $Fs(2, 28) > 5.28$, $ps_{BH} \leq 0.033$; l = 0.9 - 1 s: $Fs(2, 28) < 4.17$, $ps_{BH} > 0.052$) and for FPR at all data lengths ($Fs(2, 28) > 5.00$, $ps_{BH} < 0.05$). One-way rm-ANOVAs revealed

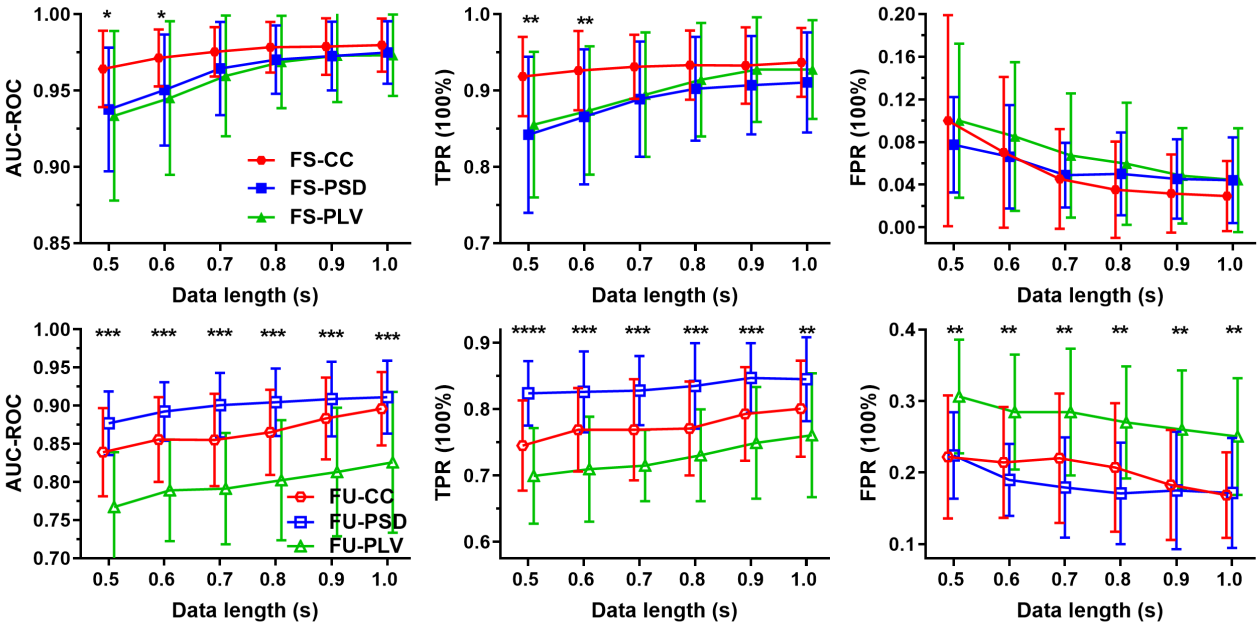

Fig. 5. Control state detection performance comparison of different algorithm frameworks (top panel: the FS framework; bottom panel: the FU framework) and different features at data length of 0.5 – 1s. The error bars indicate standard deviations. The asterisks in the subfigures indicate a significant difference between the three features obtained by one-way repeated-measure ANOVAs after Bonferroni-Holm correction (*: $p_{BH} < 0.05$; **: $p_{BH} < 0.01$; ***: $p_{BH} < 0.001$; ****: $p_{BH} < 0.0001$).

significant effects of feature for TPR of FS at short data lengths (l = 0.5s - 0.6s: $Fs(2, 28) > 8.29$, $ps_{BH} < 0.01$; l = 0.7s - 1s: $Fs(2, 28) < 5.10$, $ps_{BH} > 0.051$) and TPR of FU at all data lengths ($Fs(2, 28) > 5.68$, $ps_{BH} < 0.01$). As for FPR, the effects of feature were insignificant for FS ($Fs(2, 28) < 2.41$, $ps_{BH} > 0.65$) but significant for FU ($Fs(2, 28) > 7.11$, $ps_{BH} < 0.01$).

Post hoc analyses using paired t-tests revealed that the FS framework achieved significantly higher TPR ($ts(14) > 2.28$, $ps_{BH} < 0.05$) and lower FPR ($ts(14) < -5.27$, $ps_{BH} < 0.001$) regardless of feature and data length. Using CC features, the FS framework achieved the highest TPRs regardless of data length (l = 0.5 s: 91.8±5.2%; l = 1 s: 93.7±4.5%). The FS framework achieved the lowest FPRs with PSD features for short data lengths (l = 0.5 s: 7.7±4.5%; l = 0.6 s: 6.6±4.9%) and with CC features for long data lengths (l = 0.7 s: 4.5±4.7%; l = 0.8 s: 3.5±4.5%; l = 0.9 s: 3.2±3.7%; l = 1 s: 2.9±3.3%). The corresponding confusion matrixes of binary classification results can be found in Fig. S7 in the supplementary materials.

B. 15-Class Balanced Accuracy

Fig. 6 shows the 15-class balanced accuracies using different algorithm frameworks and features in control state detection. Two-way rm-ANOVAs revealed main effects of algrithm framework ($Fs(1, 14) > 69.71$, ps_{BH} <10e-5) and feature ($Fs(2, 28) > 5.10$, ps_{BH} <0.05), and significant interaction of these two factors ($Fs(2, 28) > 11.07$, ps_{BH} <0.001) regardless of data length. One-way rm-ANOVA showed significant effects of feature for the FU framework ($Fs(2, 28) > 6.46$, $ps_{BH} < 0.005$) at all data lengths but significant effects of feature for the FS framework were only found at short data

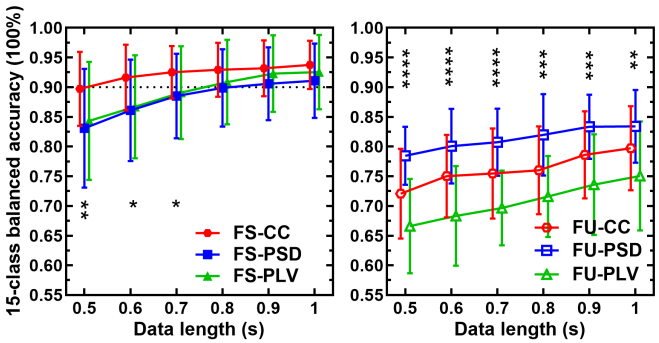

Fig. 6. Comparison of averaged 15-class balanced accuracy between different frameworks (the left panel: the FS framework; the right panel: the FU framework) and different features. The error bars indicate standard deviations. The asterisk in the subfigures indicates a significant difference between the three features obtained by one-way repeated-measure ANOVAs after Bonferroni-Holm correction (*: $p_{BH} < 0.05$; **: $p_{BH} < 0.01$; ***: $p_{BH} < 0.001$; ****: $p_{BH} < 0.0001$).

lengths (l = 0.5 s - 0.7 s: $Fs(2, 28) > 5.15$, $ps_{BH} < 0.05$; l = 0.8 s - 1 s: $Fs(2, 28) < 2.84$, $ps_{BH} > 0.16$).

Post hoc analyses using paired t-tests revealed that the FS framework achieved significantly higher balanced accuracies regardless of feature and data length ($ts(14) > 2.15$, $ps_{BH} < 0.05$). The FS framework achieved the highest balanced accuracies using CC features regardless of data length (l = 0.5s: 89.7±6.2%; l = 1 s: 93.8±4.1%). When using the PSD and PLV features, the FS frameworks also achieved excellent averaged balanced accuracies of close to or over 90% with data lengths above 0.7 s. The FU framework achieved the highest balanced accuracies using PSD features (l = 0.5 s, 78.5±4.9%; l = 1 s, 83.4±6.1%). Consistent with

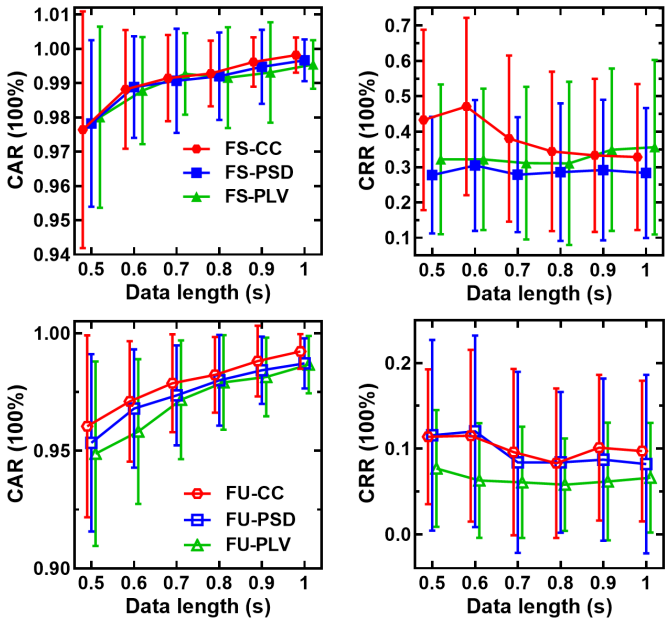

Fig. 7. Comparison of CAR and CRR for the IC data between algorithm frameworks and features. The error bars indicate standard deviations.

the balanced accuracies, the confusion matrixes (as shown in Fig. S5 and S6 in the supplementary materials) show excellent accuracies of close to or over 90% across all classes for the FS framework with CC features, even using a data length of 0.5 s.

C. Reliability of the IC Outputs

As noted in Section III-D, we defined the CAR and CRR for the IC data by simultaneously considering the results of frequency identification and control state detection. A system of high reliability should have a high CAR and a high CRR simultaneously. Fig. 7 compares these metrics for the IC data across algorithm frameworks and features. Two-way rm-ANOVAs revealed insignificant interaction between the effects of algorithm framework and feature for the two metrics at all data lengths (CAR: $Fs(2, 28) < 4.32$, $ps_{BH} > 0.13$; CRR: $Fs(2, 28) < 3.82$, $ps_{BH} > .17$) except for the CRR at data length of 0.6 ($Fs(2, 28) = 6.19$, $p_{BH} = 0.036$). Simple main effects analyses showed that the two metrics were not significantly impacted by feature (CAR: $Fs(2, 28) < 4.81$, $ps_{BH} > 0.09$; CRR: $Fs(2, 28) < 5.35$, $ps_{BH} > 0.06$) but they were significantly impacted by algorithm framework (CAR: $Fs(1, 14) > 13.63$, $ps_{BH} < 0.01$; CRR: $Fs(1, 14) > 45.12$, $ps_{BH} < 0.0001$).

Post hoc analyses showed that the FS framework had significantly higher CAR ($ts(14) > 2.58$, $ps_{BH} < 0.05$) and CRR ($ts(14) > 4.45$, $ps_{BH} < 0.001$) than the FU framework regardless of feature and data length. The FS framework achieved a higher average CAR with PLV features at short data length ($l = 0.5$ s) but with CC features at longer data length ($l > 0.6$ s). The FS framework achieved an excellent average CAR of above 97% regardless of feature and reached an average CAR of 99.8% with CC features using 1-s long data. The CRRs of the FS framework with CC features were higher at short data lengths, reaching 47.1±6.6% and 32.9±5.3% at data lengths of 0.6 s and 1 s, respectively, while the averaged CRRs of the FU framework were lower than 15% regardless of feature and data length.

V. Results of Online Experiment

Because the CC features are of lower computational cost and outperformed other features in the FS framework in the offline experiment, CC features were used for control state classification in both online systems. Table II lists the results of the main performance metrics of the online cue-guided task for FS and FU systems. Using averaged data length of 591.63±5.65 ms, the averaged SSVEP recognition accuracy, ITR, TPR, and FPR achieved by the FS system were 92.75±4.29%, 124.95±12.35 bits/min, 93.16±4.4%, and 5.21±5.85%, respectively. Using significantly longer data of 606.33±7.77 ms, these metrics achieved by the FU system significantly underperformed the FS system according to the results of paired t-tests shown in Table II. The distributions of data lengths for each subject (as shown in Fig. S8 in the supplementary materials) indicate that both the FS and FU systems could make decisions using data shorter than 700 ms, but there were more trials using data longer than 700 ms in the FU system.

The averaged 15-class confusion matrixes in Fig. 8 show higher accuracies across all classes for the FS system. For most SSVEP frequencies, the percentages of misclassified as NC state were less than 10% in the FS system but greater than 10% in the FU system. The percentage of misidentifying NC as any of the SSVEP frequencies was less than or equal to 1% in the FS system, while that ranged from 1.3% to 2.8% in the FU system. Consistent with the results of the confusion matrixes, the balanced accuracies of the FS system were significantly higher than the FU system (FS: 92.89±4.02%; FU: 79.56±6.09%; $t(13)$=10.35, p <10e-5). The FS system also achieved significantly higher CAR (FS: 99.57±0.57%; FU: 98.33±1.38%; $t(13)$=3.54, $p = 0.0036$) and CRR (FS: 29.42±27.10%; FU: 6.06±5.53%; $t(13)$=3.03, $p = 0.0097$) than the FU system, as shown in the left panels of Fig. 8. Demonstration videos showing the experimental procedures and results for two typical subjects are available online (https://www.youtube.com/watch?v=zjv4r45otNg&list=PL1uuhAxV-nu1GuLvPgbGWAjfGru81GYct).

VI. Discussions

Improving the performance of control state detection for asynchronous SSVEP-BCI is a critical issue for free communication in various practical applications. In most existing studies, it is still a challenge to accurately and quickly classify a user's control state. By employing common machine learning methods and feature extraction methods in an FS manner, this study provides a novel approach towards an efficient algorithm framework of higher accuracy with short data length in control state detection for SSVEP-BCI. The highest AUC-ROCs of 0.964±0.025 and 0.980±0.017 were obtained in offline IC/NC classification with data lengths of 0.5 s and 1 s, respectively. The FS system also achieved excellent online performance with SSVEP accuracy, ITR,

TABLE II
THE MAIN RESULTS OF THE ONLINE EXPERIMENT FOR THE FS AND FU SYSTEMS

	Accuracy of SSVEP (%)		ITR (bits/min)		TPR (%)		FPR (%)		Averaged Data Length (ms)	
	FS	FU	FS	FU	FS	FU	FS	FU	FS	FU
S1 †	**98.57**	87.14	**143.66**	107.96	**99.29**	87.14	**7.69**	39.74	**589.61**	607.56
S2	**91.43**	82.86	**119.86**	97.73	**92.14**	84.29	**8.97**	30.77	**595.83**	606.84
S3	**92.14**	80.71	**123.02**	93.89	**92.14**	82.86	**3.85**	25.64	**586.88**	599.76
S4	**95.71**	87.14	**133.35**	109.52	**97.14**	88.57	**1.28**	17.95	**591.92**	598.09
S5 †	**93.57**	70.00	**126.74**	71.07	**95.00**	72.14	**15.38**	35.90	**597.40**	605.35
S6 †	**91.43**	75.71	**121.10**	83.02	**91.43**	76.43	**2.56**	29.49	**590.34**	598.46
S7 #	**86.43**	74.29	**106.16**	78.51	**87.14**	77.86	**12.82**	29.49	**606.78**	624.17
S8 †	**89.29**	77.86	**114.92**	86.40	**89.29**	80.71	**0.00**	28.21	**585.80**	615.84
S9	**94.29**	90.00	**129.22**	115.55	**94.29**	90.71	**0.00**	21.79	**583.84**	610.33
S10	**95.00**	74.29	**131.03**	79.61	**95.71**	75.71	**15.38**	28.21	**593.03**	598.78
S11 †	**97.14**	79.29	**138.14**	89.59	**97.14**	79.29	**0.00**	34.62	**590.16**	613.27
S12	**96.43**	86.43	**135.44**	106.39	**96.43**	87.86	**0.00**	25.64	**590.08**	609.35
S13 #	**94.29**	85.00	**128.46**	103.40	**94.29**	85	**5.13**	33.33	**589.15**	601.08
S14 #	**82.86**	71.43	**98.30**	73.67	**82.86**	72.14	**0.00**	21.79	**592.12**	599.77
Mean±SD	**92.75±4.29**	80.15±6.43	**124.95±12.35**	92.59±14.44	**93.16±4.4**	81.47±6.09	**5.21±5.85**	28.75±5.98	**591.63±5.65**	606.33±7.77
$t(13)$	9.109 ****		9.988 ****		8.185 ****		14.420 ****		7.009 ****	

Note: ACC of SSVEP means the percentage of SSVEP trials correctly identified in SSVEP identification and correctly classified as IC state.
† no SSVEP-BCI experience before this study; # participated in the offline experiment; **** p<0.0001.

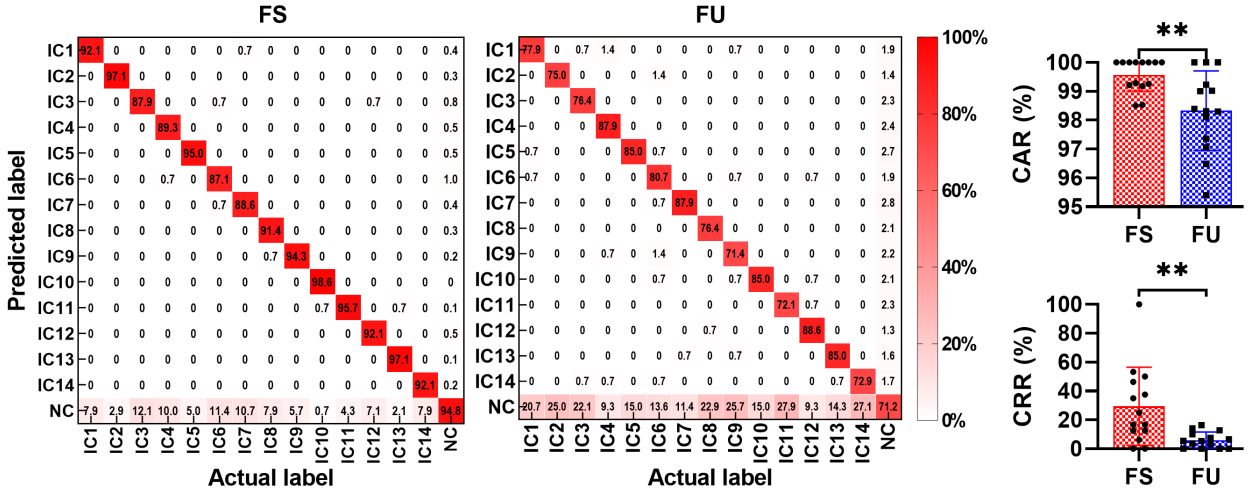

Fig. 8. Averaged 15-class confusion matrixes, averaged CARs, and CRRs of the online experiment. IC1, IC2, . . . , and IC14 correspond to SSVEPs of 8 Hz, 8.5 Hz, . . . , and 14.5 Hz, respectively.

TPR, and FPR of 92.75±4.29%, 124.95±12.35 bits/min, 93.16±4.4%, and 5.21±5.85%, respectively, using data length of 591.63±5.65 ms. It should be noted that there were three different NC states in the offline and online experiments and the FS framework still achieved low FPRs. These results suggest that high-speed and high-performance asynchronous SSVEP-BCI systems can be built using the FS framework to adapt to different working contexts. In contrast to studies using physiological signals to implement switches [13], [14], [15], [16], [17], [18], [19], [20], the method in this study requires no additional action to transfer between IC and NC states. It is a more convenient and friendly way to implement asynchronous SSVEP-BCIs. Compared with similar studies, this study substantially improves the speed of asynchronous SSVEP-BCIs by improving the recognition performance at short data lengths.

The advantage in offline and online CARs further indicates that the IC outputs of the FS framework were more reliable than those of the FU framework. Although the CRRs of the FS framework were not so high as the CARs, the results have shown an advantage of the FS framework: it can keep almost all of the correctly identified SSVEP trials and reject nearly one-third of the wrongly identified ones at the same time. For practical applications, the FS framework could reduce the negative effect of misidentified SSVEP epochs by classifying more misidentified SSVEP epochs into NC state, thus improving the reliability of an asynchronous SSVEP-BCI system.

Compared with the FU framework, the performance improvement of the FS framework can be attributed to the frequency-specific manner in feature extraction and classifier training and application since the other elements in this

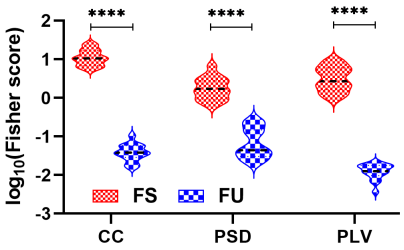

Fig. 9. Comparison of averaged Log-transformed Fisher score. The asterisks indicate significant differences obtained by paired t-tests after Bonferroni-Holm correction (****: $p_{BH} < 0.0001$).

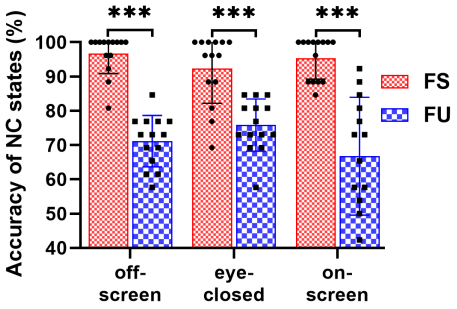

Fig. 10. Recognition accuracies of different NC states. ***$p_{BH} < 0.0001$.

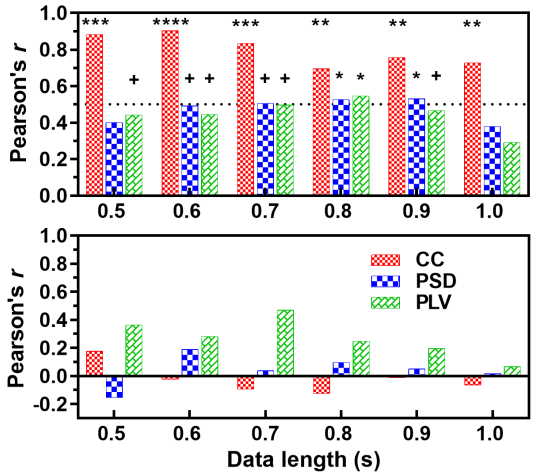

Fig. 11. Correlation between the accuracies of SSVEP identification and the AUC-ROCs of control state detection (top panel: FS framework; bottom panel: FU framework). The crosses and asterisks indicate the significance levels of the correlation coefficients (+: $p < 0.1$; *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; ****: $p < 0.0001$).

study were identical. To interpret the results, we analyzed the training data and quantified the degree of separability of the CC, PSD, and PLV features for the FS and FU frameworks separately, using the Fisher score [43]. The Fisher score can measure the class separability of a feature between IC and NC states by calculating the ratio of the between-class variance to the within-class variance, namely $\mathrm{F-score} = (m_{\mathrm{IC}} - m_{\mathrm{NC}})^2 / (\delta^2_{\mathrm{IC}} + \delta^2_{\mathrm{NC}})$, where m and δ^2 correspond to the mean value and the variance of IC or NC features, respectively. The overall separability indexes of the three kinds of features were calculated by averaging across features within each kind of features. Fig. 9 shows the comparison of the Log-transformed Fisher score. Two-way rm-ANOVA showed significant effects of algorithm framework ($F(1, 14) = 11.81$, $p < 0.0001$) and feature ($F(2, 28) = 40.05$, $p < 0.0001$) and their interaction ($F(2, 28) = 127.9$, $p = 0.0001$). Paired t-tests revealed that all three kinds of features extracted in the FS manner had significantly higher overall separability than the FU manner (CC: $t(14) = 31.36$, $ps_{BH} < 0.0001$; PSD: $t(14) = 24.23$, $ps_{BH} < 0.0001$; PLV: $t(14) = 11.36$, $ps_{BH} < 0.0001$). Considering the differences between the two frameworks in terms of feature extraction, these results indicate that the FS manner could increase the overall distance between IC and NC feature values by extracting features only related to a specific stimulus frequency and preventing contamination of feature separability from other frequencies when classifying between NC and a specific frequency. Comparisons between features found that the CC features had the highest separability for the FS framework ($ts(14) > 11.26$, $ps_{BH} < 0.0001$) while the PSD features were the best for the FU framework ($ts(14) > 2.31$, $ps_{BH} < 0.05$). These results are consistent with the control state classification performance of the offline experiment and suggest that the CC features using the FS feature extraction framework can better characterize the differences between IC and NC states in SSVEP-BCIs. A possible reason for these results is that the PSD and PLV features of single-trial EEG are more likely to be affected by noise due to the short data length. In contrast, the CC feature, a statistical feature of the relationship between the SSVEP template and the single-trial EEG epoch, is likely less affected by random noise.

Another question worth discussing is whether EEG epochs from all three NC states were effectively recognized as NC state, although the number of samples used to train the classifiers in both frameworks were balanced among different NC states. For the offline experiment results, comparisons of recognition accuracies (as shown in Fig. S9 in the supplementary materials) between different NC states found an insignificant effect for both frameworks regardless of feature and data length ($ps_{BH} > 0.05$). The online recognition accuracies of the three different NC states are shown in Fig. 10. One-way rm-ANOVAs performed to examine the effect of the type of NC state also found no significance for both online systems (FS: $F(2, 26)=2.01$, $p = 0.15$; FU: $F(2, 26)=1.91$, $p = 0.16$). These results indicate that the methods used in this study can cope with different NC states.

Although the FS framework has the advantage of obtaining high performance with short data lengths, there are also some disadvantages. The SSVEP identification result would impact the control state detection procedure in the FS framework since the SSVEP identification result determined the classifier used to recognize the control state. Therefore, we analyzed the Pearson's correlation coefficient between the AUC-ROCs of control state detection and the accuracies of SSVEP identification. As shown in Fig. 11, the results revealed positive correlations between the SSVEP accuracies and the AUC-ROCs of the FS framework, regardless of feature and data length. The correlation coefficients showed strong associations between the SSVEP accuracies and the AUC-ROCs when using CC

features (Pearson's $rs > 0.69$, $ps<0.01$). For PSD and PLV features, the correlation coefficients indicated medium associations with significance approaching marginal or greater. In contrast, the FU framework exhibited a small and insignificant association between the SSVEP accuracies and the AUC-ROCs regardless of feature and data length. The positive association between SSVEP identification performance and the control state detection performance of the FS framework may prove a double-edged sword, enhancing control state detection performance for high-performance SSVEP systems but worsening it for poor-performance ones.

Although the FS framework achieved impressive performance using common machine learning and feature extraction methods, the supervised machine learning that depends on individual training data may result in high costs in practical applications. Future research is needed to combine the FS framework with subject or stimulus transfer methods [44], [45], [46], a more efficient frequency feature extraction method [47], or the training-free method [48] to reduce the training cost. Towards high-speed and high-accuracy asynchronous SSVEP-BCIs, the performance and data length could be further optimized by incorporating more efficient dynamic stopping strategies [33], [40] into the FS framework by determining the data length adaptively for each trial. Last and most important, work is needed to validate the performance of the FS framework in more complex tasks like robot control for practical applications, as has been done in the previous studies [25], [49].

VII. Conclusion

Towards fast and accurate asynchronous SSVEP-BCI, a novel FS algorithm framework using common features and a common machine learning method was proposed and evaluated for control state detection in offline and online experiments. The offline performance using short data lengths showed that the proposed FS framework achieved excellent performance and significantly outperformed the FU framework regarding control state detection performance, balanced accuracy, and the reliability of IC outputs. In an online cue-guided selection task, the FS framework incorporating a simple dynamic stopping strategy achieved excellent performance with SSVEP accuracy, ITR, TPR, and FPR of 92.75±4.29%, 124.95±12.35 bits/min, 93.16±4.4%, and 5.21±5.85%, respectively, using an averaged data length of 591.63±5.65 ms. The reliability analysis results for the online IC outputs showed that the FS system was of high reliability with high averaged CARs of 99.57±0.57%, and almost one-third of the wrongly identified IC trials were rejected. Since the proposed FS framework can achieve high performance in control state detection with short data length, it may have great potential to be applied to improve the speed and accuracy of asynchronous SSVEP-BCIs for real-life applications.

VIII. Acknowledgment

The authors would like to thank the participants in this study.

References

[1] J. R. Wolpaw et al., "Brain–computer interface technology: A review of the first international meeting," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 2, pp. 73–164, Jun. 2000, doi: 10.1109/TRE.2000.847807.

[2] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain–computer interfaces for communication and control," *Clin. Neurophysiol.*, vol. 113, no. 6, pp. 91–767, Jun. 2002, doi: 10.1016/s1388-2457(02)00057-3.

[3] M. M. Moore, "Real-world applications for brain–computer interface technology," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, no. 2, pp. 5–162, Jun. 2003, doi: 10.1109/TNSRE.2003.814433.

[4] A. Bashashati, M. Fatourechi, R. K. Ward, and G. E. Birch, "A survey of signal processing algorithms in brain–computer interfaces based on electrical brain signals," *J. Neural Eng.*, vol. 4, no. 2, pp. R32–R57, Jun. 2007, doi: 10.1088/1741-2560/4/2/R03.

[5] X. Chen, Y. Wang, M. Nakanishi, X. Gao, T.-P. Jung, and S. Gao, "High-speed spelling with a noninvasive brain–computer interface," *Proc. Nat. Acad. Sci. USA*, vol. 112, no. 44, pp. E6058–E6067, Nov. 2015.

[6] M. Nakanishi, Y. Wang, X. Chen, Y.-T. Wang, X. Gao, and T.-P. Jung, "Enhancing detection of SSVEPs for a high-speed brain speller using task-related component analysis," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 1, pp. 104–112, Jan. 2018.

[7] Y. Zhang et al., "Correlated component analysis for enhancing the performance of SSVEP-based brain-computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 5, pp. 948–956, May 2018, doi: 10.1109/TNSRE.2018.2826541.

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

[9] A. Ravi, N. H. Beni, J. Manuel, and N. Jiang, "Comparing user-dependent and user-independent training of CNN for SSVEP BCI," *J. Neural Eng.*, vol. 17, no. 2, Apr. 2020, Art. no. 026028, doi: 10.1088/1741-2552/ab6a67.

[10] X. Zhang, S. Qiu, Y. Zhang, K. Wang, Y. Wang, and H. He, "Bidirectional Siamese correlation analysis method for enhancing the detection of SSVEPs," *J. Neural Eng.*, vol. 19, no. 4, Aug. 2022, Art. no. 046027, doi: 10.1088/1741-2552/ac823e.

[11] Y. Li, J. Xiang, and T. Kesavadas, "Convolutional correlation analysis for enhancing the performance of SSVEP-based brain–computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 12, pp. 2681–2690, Dec. 2020, doi: 10.1109/TNSRE.2020.3038718.

[12] N. Waytowich et al., "Compact convolutional neural networks for classification of asynchronous steady-state visual evoked potentials," *J. Neural Eng.*, vol. 15, no. 6, Dec. 2018, Art. no. 066031, doi: 10.1088/1741-2552/aae5d8.

[13] E. Erkan and M. Akbaba, "A study on performance increasing in SSVEP based BCI application," *Eng. Sci. Technol., Int. J.*, vol. 21, no. 3, pp. 421–427, 2018.

[14] Y. Zhu, Y. Li, J. Lu, and P. Li, "A hybrid BCI based on SSVEP and EOG for robotic arm control," *Frontiers Neurorobotics*, vol. 14, Nov. 2020, Art. no. 583641, doi: 10.3389/fnbot.2020.583641.

[15] Y. Zhou, S. He, Q. Huang, and Y. Li, "A hybrid asynchronous brain–computer interface combining SSVEP and EOG signals," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 10, pp. 2881–2892, Oct. 2020, doi: 10.1109/TBME.2020.2972747.

[16] J. Ha, S. Park, C.-H. Im, and L. Kim, "A hybrid brain–computer interface for real-life meal-assist robot control," *Sensors*, vol. 21, no. 13, p. 4578, Jul. 2021, doi: 10.3390/s21134578.

[17] X. Chai et al., "A hybrid BCI-controlled smart home system combining SSVEP and EMG for individuals with paralysis," *Biomed. Signal Process. Control*, vol. 56, Feb. 2020, Art. no. 101687.

[18] 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. 66–3156, Nov. 2013, doi: 10.1109/TBME.2013.2270283.

[19] J. Huang et al., "Hybrid asynchronous brain–computer interface for yes/no communication in patients with disorders of consciousness," *J. Neural Eng.*, vol. 18, no. 5, Apr. 2021, Art. no. 056001, doi: 10.1088/1741-2552/abf00c.

[20] C.-H. Han, E. Kim, and C.-H. Im, "Development of a brain–computer interface toggle switch with low false-positive rate using respiration-modulated photoplethysmography," *Sensors*, vol. 20, no. 2, p. 348, Jan. 2020.

[21] B. Xia, X. Li, H. Xie, W. Yang, J. Li, and L. He, "Asynchronous brain–computer interface based on steady-state visual-evoked potential," *Cognit. Comput.*, vol. 5, no. 2, pp. 243–251, 2013.

[22] P. F. Diez, V. A. Mut, E. M. Avila Perona, and E. Laciar Leber, "Asynchronous BCI control using high-frequency SSVEP," *J. Neuroengineering Rehabil.*, vol. 8, no. 1, pp. 1–9, 2011.

[23] J. Pan, Y. Li, R. Zhang, Z. Gu, and F. Li, "Discrimination between control and idle states in asynchronous SSVEP-based brain switches: A pseudo-key-based approach," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 21, no. 3, pp. 435–443, May 2013.

[24] S. Ajami, A. Mahnam, and V. Abootalebi, "An adaptive SSVEP-based brain–computer interface to compensate fatigue-induced decline of performance in practical application," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 11, pp. 2200–2209, Nov. 2018, doi: 10.1109/TNSRE.2018.2874975.

[25] L. Chen et al., "Adaptive asynchronous control system of robotic arm based on augmented reality-assisted brain–computer interface," *J. Neural Eng.*, vol. 18, no. 6, Nov. 2021, Art. no. 066005, doi: 10.1088/1741-2552/ac3044.

[26] C. Yang, X. Yan, Y. Wang, Y. Chen, H. Zhang, and X. Gao, "Spatio-temporal equalization multi-window algorithm for asynchronous SSVEP-based BCI," *J. Neural Eng.*, vol. 18, no. 4, Jul. 2021, Art. no. 0460b7, doi: 10.1088/1741-2552/ac127f.

[27] P. Poryzala and A. Materka, "Cluster analysis of CCA coefficients for robust detection of the asynchronous SSVEPs in brain–computer interfaces," *Biomed. Signal Process. Control*, vol. 10, pp. 201–208, Mar. 2014.

[28] D. Zhang, B. Huang, W. Wu, and S. Li, "An idle-state detection algorithm for SSVEP-based brain–computer interfaces using a maximum evoked response spatial filter," *Int. J. Neural Syst.*, vol. 25, no. 7, Nov. 2015, Art. no. 1550030.

[29] X. Han, K. Lin, S. Gao, and X. Gao, "A novel system of SSVEP-based human–robot coordination," *J. Neural Eng.*, vol. 16, no. 1, Feb. 2019, Art. no. 016006, doi: 10.1088/1741-2552/aae1ba.

[30] K. Suefusa and T. Tanaka, "Asynchronous brain–computer interfacing based on mixed-coded visual stimuli," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 9, pp. 2119–2129, Sep. 2018, doi: 10.1109/TBME.2017.2785412.

[31] X. Zhang et al., "A convolutional neural network for the detection of asynchronous steady state motion visual evoked potential," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 6, pp. 1303–1311, Jun. 2019, doi: 10.1109/TNSRE.2019.2914904.

[32] A. Ravi, J. Lu, S. Pearce, and N. Jiang, "Enhanced system robustness of asynchronous BCI in augmented reality using steady-state motion visual evoked potential," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 85–95, 2022.

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

[34] X. Chen, Y. Wang, S. Gao, T.-P. Jung, and X. Gao, "Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain–computer interface," *J. Neural Eng.*, vol. 12, no. 4, Aug. 2015, Art. no. 046008, doi: 10.1088/1741-2560/12/4/046008.

[35] R. Bruña, F. Maestu, and E. Pereda, "Phase locking value revisited: Teaching new tricks to an old dog," *J. Neural Eng.*, vol. 15, no. 5, Oct. 2018, Art. no. 056011, doi: 10.1088/1741-2552/aacfe4.

[36] A. Tharwat, T. Gaber, A. Ibrahim, and A. E. Hassanien, "Linear discriminant analysis: A detailed tutorial," *AI Commun.*, vol. 30, no. 2, pp. 169–190, 2017.

[37] J. Hua, Z. Xiong, J. Lowey, E. Suh, and E. R. Dougherty, "Optimal number of features as a function of sample size for various classification rules," *Bioinformatics*, vol. 21, no. 8, pp. 1509–1515, Apr. 2005.

[38] X. Chen, Y. Wang, M. Nakanishi, X. Gao, T.-P. Jung, and S. Gao, "High-speed spelling with a noninvasive brain–computer interface," *Proc. Nat. Acad. Sci. USA*, vol. 112, no. 44, pp. 6058–6067, Nov. 2015, doi: 10.1073/pnas.1508080112.

[39] D. H. Brainard, "The psychophysics toolbox," *Spatial Vis.*, vol. 10, no. 4, pp. 6–433, 1997. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/9176952

[40] E. Yin, Z. Zhou, J. Jiang, Y. Yu, and D. Hu, "A dynamically optimized SSVEP brain–computer interface (BCI) speller," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 6, pp. 1447–1456, Jun. 2014.

[41] S. Holm, "A simple sequentially rejective multiple test procedure," *Scandin. J. Statist.*, vol. 6, no. 2, pp. 65–70, 1979.

[42] D. E. Thompson et. al., "Performance measurement for brain–computer or brain-machine interfaces: A tutorial," *J. Neural Eng.*, vol. 11, no. 3, Jun. 2014, Art. no. 035001, doi: 10.1088/1741-2560/11/3/035001.

[43] L. Cao, J. Li, Y. Sun, H. Zhu, and C. Yan, "EEG-based vigilance analysis by using Fisher score and PCA algorithm," in *Proc. IEEE Int. Conf. Prog. Informat. Comput.*, vol. 1, Dec. 2010, pp. 175–179.

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

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

[46] W. Yan, Y. Wu, C. Du, and G. Xu, "An improved cross-subject spatial filter transfer method for SSVEP-based BCI," *J. Neural Eng.*, vol. 19, no. 4, Aug. 2022, Art. no. 046028.

[47] Y. Pei et al., "A tensor-based frequency features combination method for brain–computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 465–475, 2022, doi: 10.1109/TNSRE.2021.3125386.

[48] Y. Chen, C. Yang, X. Chen, Y. Wang, and X. Gao, "A novel training-free recognition method for SSVEP-based BCIs using dynamic window strategy," *J. Neural Eng.*, vol. 18, no. 3, Jun. 2021, Art. no. 036007.

[49] Y. Ke, P. Liu, X. An, X. Song, and D. Ming, "An online SSVEP-BCI system in an optical see-through augmented reality environment," *J. Neural Eng.*, vol. 17, no. 1, Feb. 2020, Art. no. 016066.