A Hybrid Method Fusing Frequency Recognition With Attention Detection to Enhance an Asynchronous Brain–Computer Interface

Jing Zhao, Ye Shi, Wenzheng Liu, Tianyi Zhou, Zheng Li, and Xiaoli Li

***Abstract*— Objective: One critical problem in controlling an asynchronous brain-computer interface (BCI) system is to discriminate between control and idle states. This paper proposes a hybrid attention detection and frequency recognition method based on weighted Dempster-Shafer theory (ADFR-DS), which integrates information of different aspects of the task from two brain regions, to enhance asynchronous control performance of a steady-state visual evoked potential (SSVEP)-based BCI system. Methods: The ADFR-DS method utilizes a hybrid architecture to process electroencephalogram (EEG) data from different channels simultaneously: an individualized frequency band based optimized complex network (IFBOCN) algorithm processes neural activity from the prefrontal area for attention detection, and an ensemble task-related component analysis (eTRCA) algorithm processes data from the occipital area for frequency recognition. The ADFR-DS method then fuses their classification results at decision level to generate the final output of the BCI system. A novel weighted Dempster-Shafer fusion method was proposed to enhance the fusion performance. This study evaluated the proposed method using a 40-target dataset recorded from 35 participants. Main results: The proposed method outperformed the eTRCA algorithm in the true positive rate (TPR), true negative rate (TNR), accuracy (ACC) and information transfer rate (ITR). Specifically, ADFR-DS improved the average ACC of eTRCA from 62.71% to 69.30%, and improved the average ITR from 184.28 bits/min to 216.89 bits/min (data length 0.3 s). Conclusion: The results suggest that the proposed ADFR-DS method can enhance asynchronous SSVEP-based BCI systems.**

***Index Terms*— Brain-computer interface, asynchronous classification, steady-state visual evoked potential, attention detection.**

Manuscript received 7 November 2022; revised 27 March 2023; accepted 2 May 2023. Date of publication 12 May 2023; date of current version 25 May 2023. This work was supported in part by the Natural Science Foundation of Hebei Province under Grant F2020203070, in part by the Scientific and Technological Innovation 2030 under Grant 2021ZD0204300, in part by the Hebei Province Science and Technology Support Plan under Grant 21372001D, in part by the Natural Science Fund for Distinguished Young Scholars of Hebei Province of China under Grant F2021203033, and in part by the Hebei Innovation Capability Improvement Plan Project under Grant 22567619H. *(Corresponding author: Jing Zhao.)*

Jing Zhao, Ye Shi, and Wenzheng Liu are with the School of Electrical Engineering and the Key Laboratory of Intelligent Rehabilitation and Neromodulation of Hebei Province, Yanshan University, Qinhuangdao 066004, China (e-mail: zhaoj0104@ysu.edu.cn; 1242318322@qq.com; wenzhengl0222@163.com).

Tianyi Zhou and Zheng Li are with the State Key Laboratory of Cognitive Neuroscience and Learning, Center for Cognition and Neuroergonomics, Beijing Normal University, Zhuhai 519087, China (e-mail: tianyi.Zhou@foxmail.com; lz@bnu.edu.cn).

Xiaoli Li is with the State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100875, China (e-mail: xiaoli@bnu.edu.cn).

Digital Object Identifier 10.1109/TNSRE.2023.3275547

I. Introduction

BRAIN-COMPUTER interface (BCI) is an emerging field which builds communication channels between humans and computer devices based on neural recordings. BCI systems have potential application in robotic systems that help people suffering from motor disabilities [1], [2], [3]. Current BCI systems typically use electroencephalogram (EEG) signals to record neural activity, due to its non-invasive and inexpensive hardware, ease of use, and acceptable temporal resolution [4], [5], [6]. There are three main approaches for EEG BCI: 1) steady-state visual evoked potential (SSVEP), 2) event-related potential (ERP), and 3) motor imagery (MI) potential [7], [8], [9]. Among them, SSVEP-based BCI is commonly used for control of robot devices because of its relatively high information transfer rate (ITR) and low training demands [10], [11].

SSVEP is a type of EEG signal reflecting neural responses in the visual cortex to visual stimuli flickering at specific frequencies. In a SSVEP-based BCI system, several visual stimuli of different frequencies, corresponding to different commands to a robot, are used to evoked the SSVEP responses [12]. To classify the user's intended command from SSVEP signals, various synchronous methods have been proposed to identify the frequency or phase components related to the flickering stimuli, such as power spectral density analysis (PSDA), multivariate synchronization index (MSI), canonical correlation analysis (CCA), and its extensions such as MsetCCA, MEMD-CCA and FBCCA [13], [14], [15], [16], [17]. Notably, Nakanishi et al. proposed an ensemble task-related component analysis (eTRCA) algorithm to extract the task-related features and improved the signal-noise ratio of SSVEP signals [18].

Compared with synchronous methods, the asynchronous methods recognize control/idle state along with the target frequency, allowing the user to start or stop the mental task (for SSVEP, fixating on the stimulus) at will, providing more flexibility and more natural control. However, as the user is allowed to perform mental activity other than fixating on stimuli during the idle state, it is difficult to model the variety of patterns which may appear in idle state neural

activity. Instead, most research efforts in state-of-the-art asynchronous methods have been devoted to modeling SSVEP responses in the control state. Waytowich et al. and Zhang et al. built convolutional neural networks to enhance the classification of control and idle states [19], [20]. Mao et al. proposed a recursive Bayesian-based approach to improve the classification efficiency [21]. Zhang et al. improved signal-noise ratio (SNR) of control state responses by proposing a maximum evoked response (MER) spatial filter [22], and Han et al. combined the SNR and phase-lock-value (PLV) features to build a maximum SNR and maximum PLV (MSMP) spatial filter [23]. Although these works reported successful asynchronous control of different robot devices, their true negative rates (TNRs) of distinguishing idle state need to be improved. TNR denotes the percentage of the non-control commands that are correctly detected and is an important criterion to evaluate the asynchronous methods. A method with lower TNR triggers more false commands when the subject is not intended to control, which may damage robots and surroundings in real-world applications.

To improve the TNR of asynchronous BCI systems, some researches adopted hybrid approaches integrating additional EEG signals to provide an ON/OFF switch. Pfurtscheller et al. implemented an MI-based brain switch to activate or deactivate an SSVEP-based orthosis [24]. Xu et al. utilized motor execution and movement-related cortical potentials as brain switches to start or stop cursor movement [25], [26]. Mao et al. fused a hybrid BCI with machine intelligence to enhance the robot executive efficiency [12]. Yousefi et al. detected presence of error-related potentials to block unintended BCI output [27]. These works, which required more time and mental burden for the users to frequently switch from one mental task to another, demonstrated significantly higher TNRs but poorer TPRs.

Our previous work found significant differences in attention levels between control and idle states during SSVEP task [28], [29]. We proposed an individualized frequency band based optimized complex network (IFBOCN) algorithm to detect the attention level of the user from a single FPz channel, and demonstrated its effectiveness in recognizing idle state for an SSVEP system [29]. The primary goal of the current work is to develop an efficient hybrid method to integrate attention detection as an auxiliary information pathway to enhance both TPR and TNR performance of asynchronous BCIs. As the attention-detection pathway is parallel to the frequency-recognition pathway, this hybrid approach requires no switching between different mental tasks and may help enhance other asynchronous control modalities. Motivated by this idea, we here propose a novel hybrid attention detection and frequency recognition method based on weighted Dempster-Shafer theory(ADFR-DS) to fuse decisions of the IFBOCN and eTRCA algorithms.

The remainder of this paper is organized as follows. Section II describes the ADFR-DS methods and the procedures for performance evaluation. Section III presents results from 35 participants. Section IV discusses our method's benefits and highlights key findings. Section V gives suggestions for future work.

II. Materials and Methods

A. Considered EEG Data

This study evaluated the performance of the proposed ADFR-DS method using a benchmark dataset acquired in an offline SSVEP-based speller experiment with 40 targets [30]. In the experiment, a user interface presents flickering 40 characters as the SSVEP stimuli. The interface was implemented using the Psychtoolbox toolbox on an LCD monitor with 60 frames per second. The stimulation frequencies are set to 8 Hz to 15.8 Hz with an interval of 0.2 Hz, respectively. The data was recorded from 35 healthy participants (18 males and 17 females) aged 17-34 with normal or correct-to-normal vision. The participants were asked to focus on a target stimulus when the stimuli were flickering. Each participant underwent six blocks of offline experiments. In each block, the 40 characters were selected as target in a random order and a total 40 trials of EEG data were recorded. Each trial lasted for 6 s, containing 0.5 s' cue guide, 5 s' flickering task, and 0.5 s' rest.

The dataset was recorded using a 64-channel Synamps2 system (Neuroscan, Inc.) at a sampling rate of 1000 Hz. The ground electrode was placed on midway between Fz and FPz, and the reference electrode was placed on the vertex. The acquired data were downsampled to 250 Hz and were preprocessed using a notch filter at 50 Hz. 11 channels, placed at FPz, Pz, P3, P4, PO3, PO4, PO7, PO8, Oz, O1, and O2 according to the International 10-20 system, were used for further analysis. Denote d as the time length used for offline analysis. The control trials were extracted between 0.64 s and 0.64+d s of the dataset, which had a delay of 0.14 s after start of stimulation considering the latency delay in human's visual system [17]. The idle trials were extracted between 0 s and 0+d s. During this period, the subject prepared for the SSVEP task but did not activate any command by staring at the corresponding stimulus. Different data lengths were selected to adequately evaluate the proposed method. As each trial of the dataset contained an idle period of 0.5 s, this paper selected four data lengths, i.e., 0.2 s, 0.3 s, 0.4 s and 0.5 s, for evaluation.

B. IFBOCN-Based Attention Detection

The IFBOCN algorithm combines complex network theory and phase space based non-linear dynamic analysis for analyzing non-linear time series [29]. To extract attention-related features, IFBOCN processes EEG signals from a single FPz channel on the prefrontal region. The prefrontal cortex is engaged in the processing of information essential for the allocation of bottom-up attention, and its EEG signal is commonly used for cognitive state estimation [31]. Our previous work also demonstrated the effectiveness of IFBOCN to detect attention using the FPz channel [29]. Starting with EEG data $\boldsymbol{x} \in \mathbb{R}^{1\times N_S}$ recorded in the h-th trial from a single channel FPz, the IFBOCN algorithm first filters $\boldsymbol{x}$ into three sub-band components $\boldsymbol{x}^{(i)}$ with three participant-individualized frequency bands. Here, N_s is the number of sampling points, and i = 1,2,3 is the index of the individualized frequency band. The individualized bands are selected for each participant from five bands of delta (0.5-3. Hz), theta (4-7 Hz), alpha (8-13 Hz), beta (13-30 Hz) and gamma (30-60 Hz). For each subject, the training data were used to compute the classification parameters and training accuracies for all the five bands, and the first three bands with highest accuracies were selected as the optimal bands for this individual. The filtered data $\boldsymbol{x}^{(i)}$=$(x_1, x_2, \cdots, x_{N_s})$ of the i-th band are then translated into the state vector $\widehat{\boldsymbol{X}}^{(i)}$ via time-delay embedding. The time-delay embedding method remodels the 1-dimension time series $\boldsymbol{x}^{(i)}$

into an M-dimension phase space [28]. The reconstruction space can be shown as follows.

$$\widehat{\boldsymbol{X}}^{(i)} = \begin{pmatrix} \boldsymbol{X}_1^{(i)} \\ \boldsymbol{X}_2^{(i)} \\ \cdots \\ \boldsymbol{X}_M^{(i)} \end{pmatrix} = \begin{pmatrix} x_1 & x_{1+\tau} & \cdots & x_{1+(m-1)\tau} \\ x_2 & x_{2+\tau} & \cdots & x_{2+(m-1)\tau} \\ \cdots & \cdots & \cdots & \cdots \\ x_M & x_{M+\tau} & \cdots & x_{N_s} \end{pmatrix} \quad (1)$$

where m is the embedding dimension, τ is the delay time, and $M = N_S - (m-1)\tau$ is the number of state vectors.

The IFBOCN method then constructs a recurrence network by using the state vectors $\boldsymbol{X}_1^{(i)}, \boldsymbol{X}_2^{(i)}, \cdots, \boldsymbol{X}_M^{(i)}$ as the nodes. To model the recurrence patterns between the nodes, IFBOCN estimates their Euclidean distances to build the edges:

$$a_{jk} = \begin{cases} 1, & \text{if } d_{jk} \le \theta \\ 0, & \text{if } d_{jk} > \theta \end{cases} \quad (2)$$

Here, θ is the threshold for connectivity between two nodes, and $d_{jk} = \sum_{n=1}^{M} \left\| \boldsymbol{X}_j(n) - \boldsymbol{X}_k(n) \right\|$ is the Euclidean distance between the j-th node and the k-th node. The parameters m, τ, and θ, which influence the network structure and detection accuracy, are optimized for each participant as in reference [29].

After constructing the recurrence network, the IFBOCN method calculates the average degree $K^{(i)}$ and the average clustering coefficient $C^{(i)}$ of the network as follows

$$K^{(i)} = \frac{1}{M} \sum_{j \in M} k_j \quad (3)$$

$$C^{(i)} = \frac{1}{N} \sum_{j \in M} C_j = \frac{1}{N} \sum_{j \in M} \frac{\sum_{k,l \in M} (w_{jk} w_{jl} w_{kl})^{1/3}}{k_j (k_j - 1)} \quad (4)$$

Here, k_j is the degree of the j^{th} node, and C_j is the clustering coefficient for the j-th node. $w_{jk} = \{0, 1\}$ is the weight between nodes j and k and is 1 if the nodes are connected. The average degrees and clustering coefficients of the three networks corresponding to the three individualized frequency bands are combined into the feature vector $\boldsymbol{F}$.

$$\boldsymbol{F} = \left[K^{(1)}, C^{(1)}, K^{(2)}, C^{(2)}, K^{(3)}, C^{(3)} \right] \quad (5)$$

The IFBOCN method then classifies the feature vector $\boldsymbol{F}$ as control or idle using a linear support vector machine (SVM) classifier. The posterior probability s_a of the SVM classifier is outputted for decision fusion.

C. eTRCA-Based Asynchronous Frequency Recognition

The eTRCA method optimizes the weight coefficients of multi-channel EEG data by maximizing the inter-trial covariance [18]. Considering the EEG data $\boldsymbol{X} \in \mathbb{R}^{N_C \times N_S}$ recorded in the h-th trial, eTRCA first decomposes EEG data into multiple sub-components $\boldsymbol{X}^{(m)}$ in a total of N_B frequency bands using filter bank analysis. Here, $N_c = 10$ denotes the number of channels for frequency recognition, N_s denotes the number of data samples, $h = 1, 2, \cdots, N_t$ indicates the trial index, $N_B = 5$ is the number of filter bands, and $m = 1, 2, \cdots, N_B$ denotes the index of sub-band component. The data $\boldsymbol{X}$ was filtered to obtain the sub-band components $\boldsymbol{X}^{(m)}$ using an array of band-pass filters. This study used zero-phase Chebyshev Type I infinite impulse response (IIR) filters, and set lower and upper cut-off frequencies to $m\times$ 8 Hz and 90 Hz, respectively.

The eTRCA method then optimizes the linear combination $\boldsymbol{v}_n^{(m)} \in \mathbb{R}^{N_C \times 1}$ and estimates the task-related component $\boldsymbol{y}^{(m)} = (\boldsymbol{v}_n^{(m)})^{\mathrm{T}} \boldsymbol{X}^{(m)}$. The optimized $\boldsymbol{y}^{(m)}$ exhibits a maximal covariance between all possible trials. eTRCA constructs the individual templates $\overline{\boldsymbol{X}}_n^{(m)}$ by averaging the training data from multiple trials. In order to extract the final feature for frequency recognition, the single-trial test data $\mathbf{X}$ and the averaged training data $\overline{\boldsymbol{X}}_n^{(m)}$ for the n-th stimulus are spatially filtered using $\boldsymbol{v}_n^{(m)}$ to remove background EEG noise. An ensemble spatial filter $\boldsymbol{V}^{(m)} = [v_1^{(m)}, v_2^{(m)}, \cdots, v_{N_f}^{(m)}]$ is used to compute the correlation coefficient between the spatially filtered data:

$$\rho_n^{(m)} = \rho \left((\boldsymbol{X}^{(m)})^T V^{(m)}, (\bar{\boldsymbol{X}}_n^{(m)})^T V^{(m)} \right) \quad (6)$$

The correlation coefficients are combined as:

$$f_n = \sum_{m=1}^{N_m} a(m) \cdot (\rho_n^{(m)})^2 \quad (7)$$

Here, N_m denotes the number of sub-bands, $n = 1, 2, \cdots, 5$ denotes the stimulus index, and $a(m) = m^{-1.25} + 0.25$ according to [18]. In this study, the number of harmonics N_h was set to 5, and the number of sub-bands N_m was set to 5. The classification output, target index T_f, is determined by the stimulus with the largest combined correlation:

$$T_f = \max_n f_n, n = 1, 2, \cdots, N_f \quad (8)$$

To distinguish control and idle states, the eTRCA-based asynchronous method trains a total of forty 2-class SVM classifiers, for classifying each stimuli frequency versus idle state. In the test stage, a single trial of testing data is processed using eTRCA. The correlation coefficient vector $[f_1, f_2, \ldots, f_n]$ is computed to recognize the target index T_f. To identify whether the testing data is in the idle state, the T_f^{th} SVM classifier is used to classify the coefficient vector into control/idle pattern. Its posterior probability s_f is outputted for decision fusion.

D. ADFR-DS Method Based on Weighted D-S Theory

The ADFR-DS method processes real-time EEG data using both the attention detection algorithm and the frequency recognition algorithm at the same time (Fig. 1). The classification decisions of the two algorithms are fused using a novel weighted D-S method to generate the final commands. The IFBOCN-based attention detection algorithm processes the EEG data recorded from the FPz channel, and classifies the brainwave pattern into 2 classes of control and idle states. The eTRCA-based asynchronous frequency recognition algorithm processes the EEG data recorded from 10 occipital channels (P3, P4, PO3, PO4, PO7, PO8, Oz, O1, and O2), and classifies the brainwave pattern into 41 classes, i.e., 40 stimulus frequencies and an idle state. As the output of the asynchronous BCIs consists of two parts, i.e., control/idle state and target frequency, the ADFR-DS method fused IFBOCN and eTRCA to generate the final decision of control/idle state, and adopted the target frequency classified by eTRCA when the brainwave pattern is in the control state.

To enhance the decision fusion performance, this study proposed a weighted D-S fusion method to assign weights to the classification probabilities of IFBOCN and eTRCA, respectively. As shown in Fig. 1, the fusion method inputs the

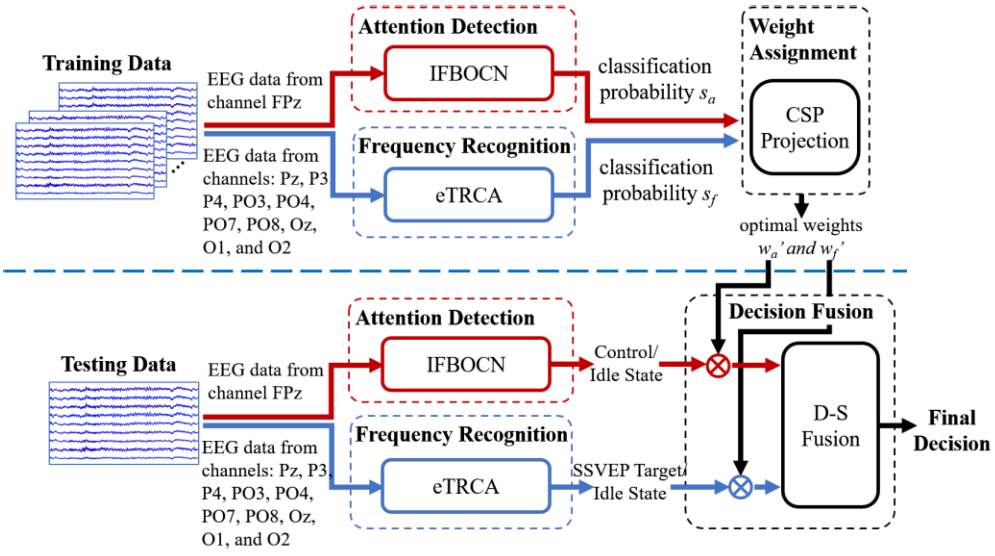

Fig. 1. Diagram of the proposed ADFR-DS method. Red path: attention detection, blue path: frequency recognition.

classification probabilities s_a of IFBOCN and s_f of eTRCA, and outputs the fused decision results based on D-S theory. The D-S theory was developed by Dempster in 1967 [32], and formalized by Shafer in 1976 [33]. Assume $\Theta = \{\theta_1, \theta_2\}$ is the set of all possible decisions, where θ_1 denotes that the participant is in the control state, and θ_2 denotes that the participant is in the idle state. $2^\Theta = \{\emptyset, \theta_1, \theta_2, \ \{\theta_1, \theta_2\}\}$ includes all the possible subsets of Θ. For a particular hypothesis A in 2^Θ, the IFBOCN and eTRCA algorithms are two sources of evidence, and are assigned the basic probability assignment (BPA) functions of $m_a(A)$ and $m_f(A)$, respectively. The weighted D-S method outputs the final decisions by fusing the BPAs of IFBOCN and eTRCA [34]:

$$m(A) = m_a(A) \oplus m_f(A) \tag{9}$$

Existing studies usually set the BPA functions $m_a(A)$ and $m_f(A)$ as the classification probabilities s_a and s_f. But as IFBOCN and eTRCA deliver varied accuracies in recognizing control and idle states, assigning weights for them can improve the overall performance. Therefore, this study used a CSP algorithm to calculate optimal weights for the two BPA functions. CSP is first introduced in EEG analysis by Koles et al. [35], and is used to compute a spatial filter to maximize the difference of variance between the two classes [36]. In this study, CSP algorithm is applied to calculate the optimal weights $\hat{\boldsymbol{w}} = [w'_a, \ w'_f]$ for transforming the probabilities $[s_a, s_f]$ to the BPA functions $m_a(A)$ and $m_f(A)$. The BPAs are calculated following a 4-step process.

(1) The probability matrix $\boldsymbol{S}$ were extracted from the control and idle trials of the training data. The CSP algorithm computes a projection matrix $\boldsymbol{W}$ to maximize the difference of variance between $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$. Here, $\boldsymbol{S}_1$ is the probability matrix of control trials, and $\boldsymbol{S}_2$ is the that of idle trials. The following eigenvalue decomposition problem is solved to obtain the projection matrix $\boldsymbol{W}$.

$$\boldsymbol{C}_1 \times \boldsymbol{W} = (\boldsymbol{C}_1 + \boldsymbol{C}_2) \times \boldsymbol{W} \times \boldsymbol{D} \tag{10}$$

Here, $\boldsymbol{C}_1$ and $\boldsymbol{C}_2$ denote the covariance matrices between the probability matrix and the corresponding control/idle pattern, and $\boldsymbol{D}$ is the diagonal matrix containing the eigenvalues of $\boldsymbol{C}_1$.

(2) As each row of the obtained $\boldsymbol{W} \in \mathbb{R}^{2\times 2}$ represents a projection vector, this study selects the optimal row by considering absolute values of all the elements in $\boldsymbol{W}$ along with training accuracies of IFBOCN and eTRCA. Assume $w_{i,j}$ is the projection element in the i-th row and j-th column of $\boldsymbol{W}$. Considering that the algorithm with better training performance should be assigned with a bigger weight, the optimal row $\boldsymbol{w}' = [w_1', w_2']$ is selected according to the following rules:

(a) If $|w_{1,1}| \geq |w_{1,2}|$ and $|w_{2,1}| < |w_{2,2}|$,

$$w' = [w'_1, w'_2] = \begin{cases} \left[|w_{1,1}|, |w_{1,2}|\right], & \text{if } (acc_a \geq acc_f) \\ \left[|w_{2,1}|, |w_{2,2}|\right], & \text{if } (acc_a < acc_f) \end{cases} \tag{11}$$

Here, acc_a denotes the training accuracy of IFBOCN, and acc_f denotes the training accuracy of eTRCA.

(b) If $|w_{1,1}| < |w_{1,2}|$ and $|w_{2,1}| \geq |w_{2,2}|$,

$$w' = [w'_1, w'_2] = \begin{cases} \left[|w_{2,1}|, |w_{2,2}|\right], & \text{if } (acc_a \geq acc_f) \\ \left[|w_{1,1}|, |w_{1,2}|\right], & \text{if } (acc_a < acc_f) \end{cases} \tag{12}$$

(c) Otherwise, if both $|w_{1,1}|$ and $|w_{2,1}|$ are bigger/smaller in the corresponding projection vector, the IFBOCN algorithm would be assigned a bigger/smaller weight. The optimal vector is selected to minimize the difference between the two weights.

$$w' = [w'_1, w'_2] = \begin{cases} \left[|w_{2,1}|, |w_{2,2}|\right], & \text{if } \left(\dfrac{|w_{1,1}|}{|w_{1,2}|} \geq \dfrac{|w_{2,1}|}{|w_{2,2}|}\right) \\ \left[|w_{1,1}|, |w_{1,2}|\right], & \text{if } \left(\dfrac{|w_{1,1}|}{|w_{1,2}|} < \dfrac{|w_{2,1}|}{|w_{2,2}|}\right) \end{cases} \tag{13}$$

(3) The weight vector is normalized as follows.

$$\hat{w} = [w'_a, w'_f] = \begin{cases} \left[1, \dfrac{w'_2}{w'_1}\right], & \text{if } (w'_1 \geq w'_2) \\ \left[\dfrac{w'_1}{w'_2}, 1\right], & \text{if } (w'_1 < w'_2) \end{cases} \tag{14}$$

(4) The weighted probabilities are assigned as the BPAs of IFBOCN and eTRCA, respectively. Considering that the classification decision is control state when s_a is bigger than 0.5, this study computes the BPA of IFBOCN as follows.

$$m_a(\theta_1) = \begin{cases} w'_a \times (s_a - 0.5) + 0.5, & \text{if } (s_a \geq 0.5) \\ w'_a \times s_a, & \text{if } (s_a < 0.5) \end{cases} \tag{15}$$

$$m_a(\theta_2) = \begin{cases} w'_a \times \overline{s_a}, & \text{if } (s_a \geq 0.5) \\ w'_a \times (\overline{s_a} - 0.5) + 0.5, & \text{if } (s_a < 0.5) \end{cases} \tag{16}$$

$$m_a(\{\theta_1, \theta_2\}) = 1 - m_a(\theta_1) - m_a(\theta_2) \tag{17}$$

Here, $\overline{s_a} = 1 - s_a$ denotes the probability of idle state given by IFBOCN, and $\{\theta_1, \theta_2\}$ denotes an uncertain hypothesis about whether the participant is in the control or idle state. The BPA of eTRCA is calculated as follows.

$$m_f(\theta_1) = \begin{cases} w'_f \times (s_f - 0.5) + 0.5, & \text{if } (s_f \geq 0.5) \\ w'_f \times s_f, & \text{if } (s_f < 0.5) \end{cases} \tag{18}$$

$$m_f(\theta_2) = \begin{cases} w'_f \times \overline{s_f}, & \text{if } (s_f \geq 0.5) \\ w'_f \times (\overline{s_f} - 0.5) + 0.5, & \text{if } (s_f < 0.5) \end{cases} \tag{19}$$

$$m_f(\{\theta_1, \theta_2\}) = 1 - m_f(\theta_1) - m_f(\theta_2) \tag{20}$$

The BPAs of both algorithms are fused based on the D-S theory to generate the final decision.

$$\begin{aligned} m_{total}(A) &= m_a(A) \oplus m_f(A) \\ &= \frac{1}{1-K} \sum\nolimits_{B \cap C = A} m_a(B) \cdot m_f(C) \end{aligned} \tag{21}$$

$$K = \sum\nolimits_{B \cap C = \emptyset} m_a(B) \cdot m_f(C) \tag{22}$$

Here, K is the degree of conflict between m_a and m_f.

E. Evaluation Procedure

To evaluate the proposed ADFR method, this study conducted an offline evaluation using the SSVEP dataset recorded from 35 participants. The recorded data were processed in a leave-one-block-out cross validation to evaluate offline performance, i.e., 6 folds of evaluation were conducted. In each fold, 1 block of control-session data and 1 block of idle-session data were selected for testing, and the remaining 10 blocks were used for training. To measure performance, we calculated classification accuracy (ACC), true positive rate (TPR), TNR, and ITR, for input EEG data length (per trial) ranging from 0.2 s to 0.5 s with an interval of 0.1 s [37]. ACC is the percentage of correctly classified trials in both control and idle sessions. TPR is the percentage of control trials in which the desired commands are successfully activated. TNR is the percentage of idle trials in which no commands are outputted. The ITR, in bits/min, is defined as follows:

$$\text{ITR} = \frac{60}{T}\left(\log_2 N + P \cdot \log_2 P + (1-P) \cdot \log_2 \frac{1-P}{N-1}\right) \tag{23}$$

where N denotes the number of classes ($N = 41$ for the asynchronous task in this study), P is the classification accuracy, and T is the time required to output a command. For offline analysis, a 0.5 s gaze-shifting time was added to the ITR calculation.

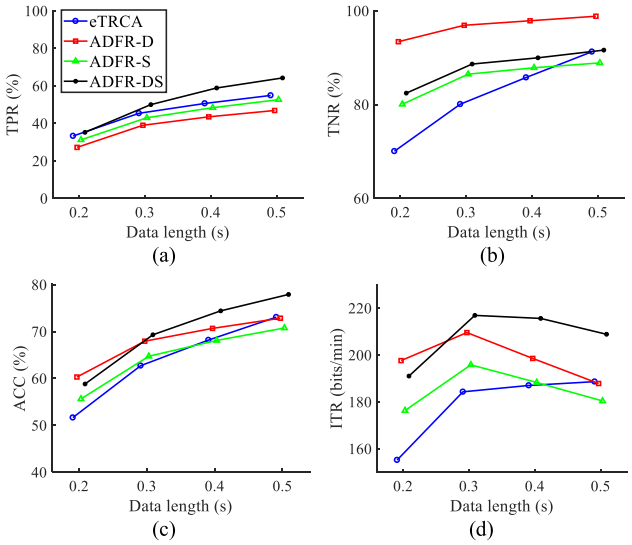

Fig. 2. Offline performance of ADFR-DS, fusing IFBOCN and eTRCA, compared to eTRCA alone and other fusion methods. (a) TPR, (b) TNR, (c) ACC, and (d) ITR.

In this study, we also implemented two decision-level fusion methods for comparison, namely ADFR with decision fusion (ADFR-D) and ADFR with score fusion (ADFR-s). Given the classification result T_a of attention detection and the result T_f of frequency recognition, ADFR-D uses T_a as an on/off switch for outputting T_f. The fused decision T is:

$$T = \begin{cases} T_f, & \text{if } (T_a = 1) \\ 0, & \text{if } (T_a = 0) \end{cases} \tag{24}$$

Here, $T_a = 0$ indicates that the decision of IFBOCN is idle state, $T_a = 1$ indicates that the decision of IFBOCN is control state, $T_f = 0$ indicates that the decision of eTRCA is idle state, and $T_f = 1,2,\ldots,40$ indicates that the decision of eTRCA is the corresponding target frequency.

The ADFR-S method fuses decisions according to the probabilities s_a and s_f calculated by the corresponding SVM classifier. It is designed to select a decision which is more likely to be correct by comparing probabilities when results from the two algorithms are in conflict. The ADFR-S method integrates the decisions according to s_a and s_f to obtain final command T:

$$T = \begin{cases} 0, & \text{if } (T_a = 0 \text{ and } T_f{=}0) \\ T_f, & \text{if } (T_a = 1 \text{ and } T_f \neq 0) \\ 0, & \text{if } (T_a = 0 \text{ and } T_f \neq 0 \text{ and } \overline{s_a} \geq s_f) \\ T_f, & \text{if } (T_a = 0 \text{ and } T_f \neq 0 \text{ and } \overline{s_a} < s_f) \\ \max\limits_n f_n, & \text{if } (T_a = 1 \text{ and } T_f = 0 \text{ and } s_a \geq \overline{s_f}) \\ 0, & \text{if } (T_a = 1 \text{ and } T_f = 0 \text{ and } s_a < \overline{s_f}) \end{cases} \tag{25}$$

Here, $\overline{s_a} = 1 - s_a$ denotes the probability of idle state given by IFBOCN, and $\overline{s_f} = \max(1 - f_n)$ denotes the probability of idle state given by eTRCA.

III. Results

Fig. 2 shows the offline classification performance of ADFR-D (red curves), ADFR-S (blue curves) and ADFR-DS

(black curves) when fusing IFBOCN and eTRCA, as compared to eTRCA alone (blue curves). Performance was evaluated at data lengths ranging from 0.2 s to 0.5 s with an interval of 0.1 s. The results indicated that the proposed ADFR-DS method achieved higher average TPRs, TNRs, ACCs, and ITRs than eTRCA, regardless of data length. As shown in Fig. 2, ADFR-DS obtained average ACCs of 58.81±11.54 % with data length of 0.2 s, 69.30±12.85 % with data length of 0.3 s, 74.40±12.33 % with data length of 0.4 s, and 77.90±11.99 % with data length of 0.5 s. The results of paired t-test showed that the accuracy of ADFR-DS was significantly higher than that of eTRCA regardless of data length ($p<0.001$). When the data length was 0.3 s, ADFR-DS achieved the highest ITR of 216.89±63.43 bits/min. Among the 4 methods, ADFR-D delivered the highest TNRs regardless of data length. Especially, ADFR-D achieved an average TNR of 98.83±1.27 % with data length of 0.5 s. But the TPRs of both ADFR-D and ADFR-S were lower than that of eTRCA alone. Compared with ADFR-D, ADFR-DS delivered higher TPRs and lower TNRs regardless of data length, and achieved significantly higher ACCs and ITRs at data length between 0.3 s and 0.5 s ($p<0.05$). When the data length was 0.2 s, ADFR-DS obtained a lower average ACC of 58.81% than ADFR-D (60.29 %). As ADFR-DS fuses decisions of eTRCA and IFBOCN for recognizing the control/idle state, it cannot correct the false decisions of frequency recognition of eTRCA. Therefore, ADFR-DS enhanced the TNR of eTRCA from 70.04 % to 82.43 % by blocking its false triggers, but achieved lower improvement of the TPR from 33.24 % to 35.19 % with data length of 0.2 s. Both TPR and TNR performance of ADFR-DS were improved with increased data length.

Fig. 3 shows the classification accuracies for the 35 participants, comparing eTRCA and ADFR-DS at data length of 0.3 s. ADFR-DS, which integrates eTRCA with IFBOCN-based attention detection, improved the average ACC of eTRCA from 62.71% to 69.30% and improved the average ITR from 184.28 bits/min to 216.89 bits/min. Notably, the participant S3 achieved an average ACC of 95.83 %, which was 17.8 % higher than that of eTRCA (80.00 %). For most of the 35 participants, ADFR-DS delivered higher ACCs than eTRCA. This demonstrates the effectiveness of the proposed method in improving asynchronous BCI system performance. In addition, IFBOCN alone achieved an average ACC of 81.36 %. It should be noted that both eTRCA and ADFR-DS are 41-class classification algorithms while IFBOCN is a 2-class algorithm only for idle state detection. Therefore, the ACC results of IFBOCN were not presented in Fig. 3.

Fig. 4 shows the TNRs for the 35 participants, comparing eTRCA, IFBOCN and ADFR-DS at data length of 0.3 s. The eTRCA algorithm alone obtained an average TNR of 80.07±8.43 %, and the IFBOCN alone obtained an average TNR of 81.11±11.05 %. The proposed ADFR-DS enhanced average TNR to 88.63±6.23 % by considering the outputs of both eTRCA and IFBOCN.

IV. Discussion

This paper proposed a novel ADFR-DS method for asynchronous control of SSVEP-based BCI systems. The ADFR-DS method uses a hybrid architecture which combines IFBOCN-based attention detection and eTRCA-based asynchronous frequency recognition. The design of the proposed method is different from previous asynchronous BCIs in three major aspects. First, as the uncertainty of BCI commands is influenced by the user's attentional state, the proposed method detects attention level in real time and combines it with frequency recognition results to decide the final output of the BCI system. Offline evaluations show that the proposed method is effective in enhancing asynchronous performance. Second, contrary to previous hybrid BCIs that use multiple mental tasks, the proposed ADFR-DS method only requires the user to conduct a single mental activity, fixate on the stimulus target, to control the BCI system. The two modules of attention detection and frequency recognition are used to extract different metrics of brain status in the single mental task. Compared with previous hybrid BCIs, the ADFR-DS method improves the accuracy of discriminating control and idle states while avoiding the time cost and mental burden of switching between different tasks. Third, the proposed method utilizes a novel IFBOCN algorithm, which has been demonstrated to outperform state-of-the-art algorithms for attention detection, to distinguish control and idle states [4], [29]. The offline results in Fig. 4 show the effectiveness of IFBOCN in separating control and idle states during the asynchronous SSVEP task. Notably, participant S3 achieved an ACC of 99.38 % using IFBOCN with data length of 0.5 s.

The proposed ADFR-DS method can use a frequency recognition algorithm of one's choosing. In this study, we implemented and evaluated a commonly-used algorithm, eTRCA. The eTRCA algorithm is first introduced in SSVEP analysis for synchronous frequency recognition [18]. It was used in previous works to recognize the target frequency rather than detect the idle state. To develop an eTRCA-based asynchronous algorithm, this study trains forty 2-class SVM classifiers to separate each stimuli frequency versus idle state. The testing data is processed using eTRCA to classify the target frequency T_f, and is then classified into control or idle state using the corresponding T_f^{th} SVM classifier. The offline results showed that the eTRCA-based asynchronous algorithm delivered average TNRs of 70.04±7.99 % with data length of 0.2 s, 80.07±8.43 % with data length of 0.3 s, 85.76±8.13 % with data length of 0.4 s, and 91.26±6.11 % with data length of 0.5 s.

As the eTRCA algorithm can be used for both frequency recognition and idle-state detection, IFBOCN provides overlapping information on idle-state detection. Therefore, this study compared the proposed method with two different methods for fusing eTRCA and IFBOCN, ADFR-D to utilize the decision of FBOCN as an on/off switch, and ADFR-S to fuse the two outputs by simply comparing their probabilities. As shown in Fig. 2, both ADFR-D and ADFR-S obtained relatively better TNR performance than eTRCA, but delivered significantly worse TPR performance. This indicates that the incorrect decisions by IFBOCN will block the output of correct decisions by eTRCA. However, ADFR-DS based on the weighted D-S method outperformed eTRCA and IFBOCN in both TPR and TNR. These results show the effectiveness of the proposed ADFR-DS methods for enhancing the asynchronous BCI performance.

Fig. 5 shows the computational time of the proposed ADFR-DS method in Matlab (MathWorks Inc., USA) on a notebook computer with the configuration of Inter(R) Core(TM) i7-7700HQ CPU @2.80 GHz, 8 GB RAM, 64-bit Win 10. The time was recorded from the end of classification process of IFBOCN and eTRCA to the end of the training or testing period. The training time and testing time were

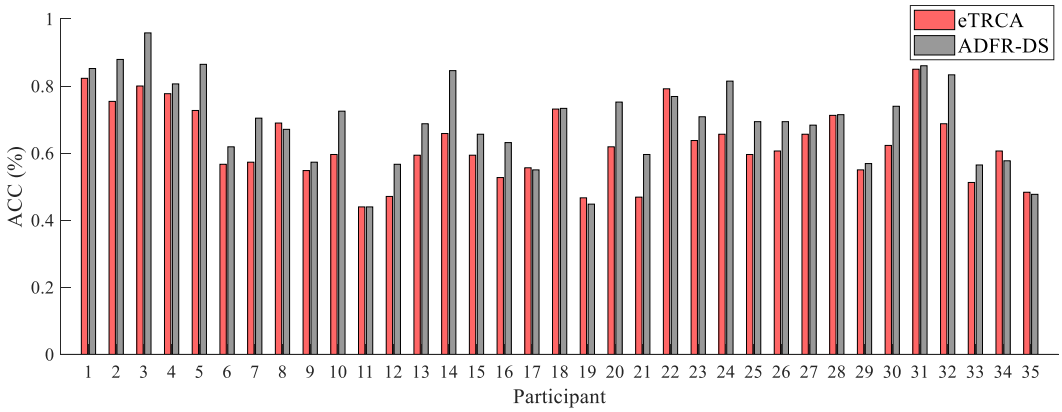

Fig. 3. ACCs of eTRCA and ADFR-DS for the 35 participants.

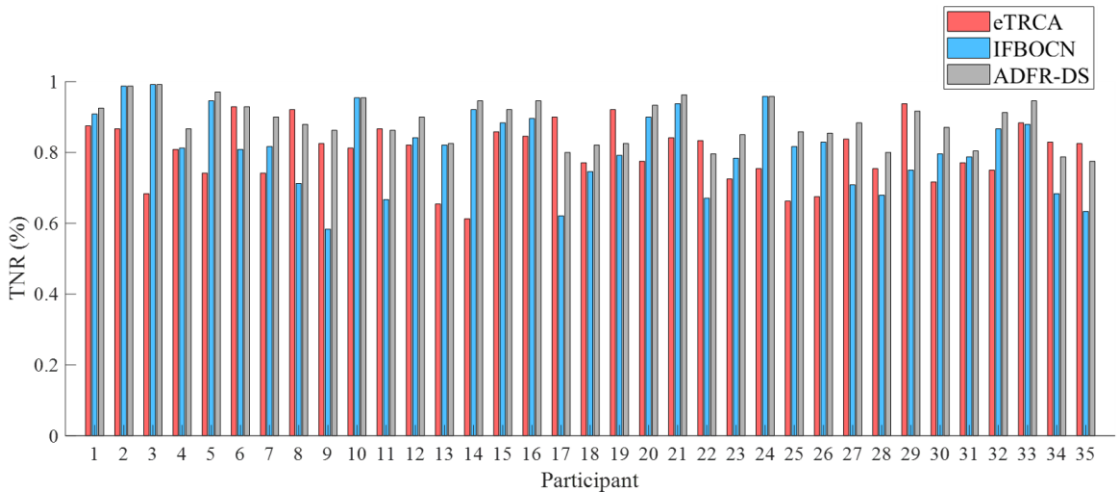

Fig. 4. TNRs of eTRCA, IFBOCN and ADFR-DS for the 35 participants.

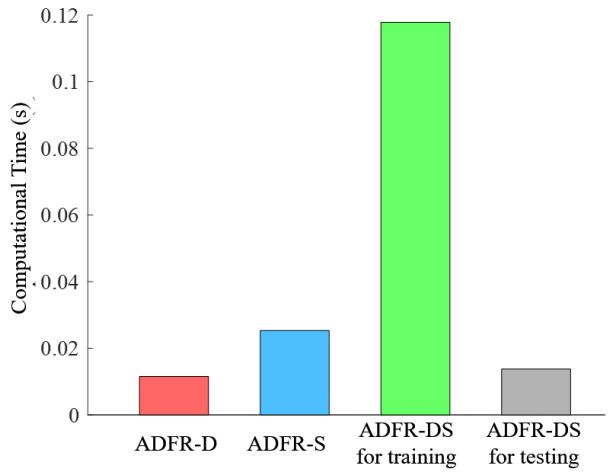

Fig. 5. Computational time of ADFR-DS compared with other fusion methods.

computed separately for ADFR-DS. As ADFR-D and ADFR-S fuses decisions with no training, Fig. 5 only presents their testing time for comparison. The experimental results showed that ADFR-D took the shortest computational time of 0.0115 s, and ADFR-S consumed the longest time of 0.0253 s for testing. ADFR-DS took 0.1178 s for training and 0.0138 s for testing. As the training step was conducted only once before the experiment, the three fusion methods consuming less than 0.05 s on computing are suitable for online applications.

V. Conclusion

The proposed ADFR-DS method enhanced TPR, TNR, ACC, and ITR, outperforming eTRCA in asynchronous classification, making it a promising new approach for asynchronous control of robotic systems, such as mobile robots and wheelchairs. In our future work, the ADFR-DS method will be further improved before application to closed-loop control of a real-world BCI system. First, more asynchronous algorithms will be fused with eTRCA and IFBOCN to enhance the classification performance. We will also try integrating a synchronous frequency recognition algorithm to improve the TPR performance [38]. Second, the decision fusion module will be enhanced by extracting more features, such as Euclidean distance and fuzzy membership, to better estimate the uncertainties of different decisions [34], [39].

References

[1] C. J. Bell, P. Shenoy, R. Chalodhorn, and R. P. N. Rao, "Control of a humanoid robot by a noninvasive brain–computer interface in humans," *J. Neural Eng.*, vol. 5, no. 2, pp. 214–220, 2008.

[2] R. Fu, H. Wang, M. Han, D. Han, and J. Sun, "Scaling analysis of phase fluctuations of brain networks in dynamic constrained object manipulation," *Int. J. Neural Syst.*, vol. 30, no. 2, Feb. 2020, Art. no. 2050002.

[3] J. Jin, Z. Chen, R. Xu, Y. Miao, X. Wang, and T. 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.

[4] Z. Liang, X. Wang, J. Zhao, and X. Li, "Comparative study of attention-related features on attention monitoring systems with a single EEG channel," *J. Neurosci. Methods*, vol. 382, Dec. 2022, Art. no. 109711.

[5] B. J. Edelman et al., "Noninvasive neuroimaging enhances continuous neural tracking for robotic device control," *Sci. Robot.*, vol. 4, no. 31, Jun. 2019, Art. no. eaaw6844.

[6] Y. Jiao et al., "Sparse group representation model for motor imagery EEG classification," *IEEE J. Biomed. Health Informat.*, vol. 23, no. 2, pp. 631–641, Mar. 2019.

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

[8] E. A. Aydin, O. F. Bay, and I. Guler, "P300-based asynchronous brain computer interface for environmental control system," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 3, pp. 653–663, May 2018.

[9] J. Jin, Y. Miao, I. Daly, C. Zuo, D. Hu, and A. Cichocki, "Correlation-based channel selection and regularized feature optimization for MI-based BCI," *Neural Netw.*, vol. 118, pp. 262–270, Oct. 2019.

[10] J. Zhao, W. Li, and M. Li, "Comparative study of SSVEP- and P300-based models for the telepresence control of humanoid robots," *PLoS ONE*, vol. 10, no. 11, pp. 1–18, Nov. 2015.

[11] 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.

[12] X. Mao, W. Li, C. Lei, J. Jin, F. Duan, and S. Chen, "A brain–robot interaction system by fusing human and machine intelligence," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 3, pp. 533–542, Mar. 2019.

[13] Y. Zhang, P. Xu, K. Cheng, and D. Yao, "Multivariate synchronization index for frequency recognition of SSVEP-based brain–computer interface," *J. Neurosci. Methods*, vol. 221, pp. 32–40, Jan. 2014.

[14] Z. Lin, C. Zhang, W. Wu, and X. Gao, "Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 2, pp. 2610–2614, Nov. 2006.

[15] Y. Zhang, G. Zhou, J. Jin, X. Wang, and A. Cichocki, "Frequency recognition in SSVEP-based BCI using multiset canonical correlation analysis," *Int. J. Neural Syst.*, vol. 24, no. 4, pp. 1450013-1–1450013-14, 2013.

[16] V. P. Oikonomou, S. Nikolopoulos, and I. Kompatsiaris, "A Bayesian multiple kernel learning algorithm for SSVEP BCI detection," *IEEE J. Biomed. Health Informat.*, vol. 23, no. 5, pp. 1990–2001, Sep. 2019.

[17] 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.

[18] 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.

[19] 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.

[20] 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.

[21] X. Mao, W. Li, H. Hu, J. Jin, and G. Chen, "Improve the classification efficiency of high-frequency phase-tagged SSVEP by a recursive Bayesian-based approach," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 3, pp. 561–572, Mar. 2020.

[22] 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.

[23] 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.

[24] G. Pfurtscheller et al., "Self-paced operation of an SSVEP-based orthosis with and without an imagery-based 'brain switch': A feasibility study towards a hybrid BCI," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 18, no. 4, pp. 409–414, Feb. 2010.

[25] R. Xu et al., "Continuous 2D control via state-machine triggered by endogenous sensory discrimination and a fast brain switch," *J. Neural Eng.*, vol. 16, no. 5, Jul. 2019, Art. no. 056001.

[26] R. Xu et al., "Endogenous sensory discrimination and selection by a fast brain switch for a high transfer rate brain–computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 8, pp. 901–910, Aug. 2016.

[27] R. Yousefi, S. Rezazadeh, and C. T. Alborz, "Development of a robust asynchronous brain-switch using ErrP-based error correction," *J. Neural Eng.*, vol. 16, no. 6, 2019, Art. no. 066042.

[28] Z.-P. Wu, W. Zhang, J. Zhao, C. Chen, and P. Ji, "Optimized complex network method (OCNM) for improving accuracy of measuring human attention in single-electrode neurofeedback system," *Comput. Intell. Neurosci.*, vol. 2019, Mar. 2019, Art. no. 2167871.

[29] 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.

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

[31] F. Katsuki and C. Constantinidis, "Bottom-up and top-down attention: Different processes and overlapping neural systems," *Neuroscientist*, vol. 20, no. 5, pp. 509–521, Oct. 2014.

[32] A. Dempster, "Upper and lower probabilities induced by a multi valued mapping," *Ann. Math. Statist.*, vol. 38, pp. 325–339, Apr. 1967.

[33] G. Shafer, *Mathematical Theory of Evidence*. Princeton, NJ, USA: Princeton Univ. Press, 1976.

[34] Y.-T. Liu, N. R. Pal, A. R. Marathe, and C.-T. Lin, "Weighted fuzzy Dempster–Shafer framework for multimodal information integration," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 1, pp. 338–352, Feb. 2018.

[35] Z. J. Koles, M. S. Lazar, and S. Z. Zhou, "Spatial patterns underlying population differences in the background EEG," *Brain Topogr.*, vol. 2, no. 4, pp. 275–284, 1990.

[36] H. Sun, J. Jin, R. Xu, and A. Cichocki, "Feature selection combining filter and wrapper methods for motor-imagery based brain–computer interfaces," *Int. J. Neural Syst.*, vol. 31, no. 09, Sep. 2021, Art. no. 2150040.

[37] J. Zhao et al., "Decision-making selector (DMS) for integrating CCA-based methods to improve performance of SSVEP-based BCIs," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 5, pp. 1128–1137, May 2020.

[38] 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.

[39] J. Zhao, W. Li, X. Mao, H. Hu, L. Niu, and G. Chen, "Behavior-based SSVEP hierarchical architecture for telepresence control of humanoid robot to achieve full-body movement," *IEEE Trans. Cognit. Develop. Syst.*, vol. 9, no. 2, pp. 197–209, Jun. 2017.