A Novel 3D Paradigm for Target Expansion of Augmented Reality SSVEP

Beining Cao, Charlie Li-Ting Tsai, Nan Zhou, Thomas Do, and Chin-Teng Lin, *Fellow, IEEE*

***Abstract*— Steady-State Visual Evoked Potentials (SSVEP) have proven to be practical in Brain-Computer Interfaces (BCI), particularly when integrated with augmented reality (AR) for real-world application. However, unlike conventional computer screen-based SSVEP (CS-SSVEP), which benefits from stable experimental environments, AR-based SSVEP (AR-SSVEP) systems are susceptible to the interference of real-world environment and device instability. Particularly, the performance of AR-SSVEP significantly declines as the target frequency increases. Therefore, our study introduced a 3D paradigm that combines flicker frequency with rotation patterns as stimuli, enabling expansion of target sets without additional frequencies. In the proposed design, in addition to the conventional frequency-based SSVEP feature, bio-marker elicited by visual perception of rotation was investigated. An experimental comparison between this novel 3D paradigm and a traditional 2D approach, which increases targets by adding frequencies, reveals significant advantages. The 12-class 3D paradigm achieved an accuracy of 76.5% and an information transfer rate (ITR) of 70.42 bits/min using 1-second EEG segments. In contrast, the 2D paradigm exhibited a lower performance with 72.07% accuracy and 62.28 bits/min ITR. The result underscores the 3D paradigm's superiority in enhancing the practical applications of SSVEP-based BCIs in AR settings, especially with shorter time windows, by effectively expanding target recognition without compromising system efficiency.**

***Index Terms*— Brain-computer interface, augmented reality, SSVEP, targets expansion, multiple bio-markers.**

Received 2 December 2024; revised 19 March 2025; accepted 15 April 2025. Date of publication 18 April 2025; date of current version 1 May 2025. This work was supported in part by the Australian Research Council (ARC) under Grant DP220100803 and Grant DP250103612, in part by ITRH under Grant IH240100016, in part by the Australian National Health and Medical Research Council (NHMRC) Ideas under Grant APP2021183, in part by the UTS Human-Centric AI Centre by GrapheneX (2023-2031), in part by the Australia Defence Innovation Hub under Grant P18-650825, and in part by the Australian Defence Science Technology Group (DSTG) under Grant 12549. *(Corresponding author: Thomas Do.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the University of Technology Sydney's Ethical Committee under Application No. ETH20-5371.

The authors are with the Computational Intelligence and Brain-Computer Interfaces Laboratory, Australian AI Institute, School of Computer Science, Faculty of Engineering and Information Technology, University of Technology Sydney, Ultimo, NSW 2007, Australia (e-mail: thomas.do@uts.edu.au).

This article has supplementary downloadable material available at https://doi.org/10.1109/TNSRE.2025.3562217, provided by the authors.

Digital Object Identifier 10.1109/TNSRE.2025.3562217

I. Introduction

BRAIN-COMPUTER interface (BCI) represents a direct communication pathway that captures human intentions without the involvement of the peripheral nervous system or muscle tissue [1]. Among the various types of BCI signals, EEG is particularly favored for its portability and user-friendliness. Several EEG-based BCI paradigms have been proposed, including steady-state visually evoked potentials (SSVEP) [2], P300 [3] and motor imagery (MI) [4]. SSVEP is a neural response to visual stimulation with specific frequencies [5]. When people observe a stimulus flickering at a constant frequency, EEG signals at that frequency and its harmonics are synchronously elicited in the visual cortex. This synchronization allows users to effectively express their intentions by focusing on the target stimulus. Due to its robust neural mechanisms and high accuracy, SSVEP-based BCI has been widely used in a variety of applications, such as keyboard operation [6] and navigation [7].

The SSVEP paradigm requires the presentation of visual stimuli at specific frequencies. Thus, the signal quality is significantly influenced by the device used to display these stimuli. Currently, most SSVEP stimuli are presented using devices such as liquid-crystal displays (LCD) [8] light-emitting diodes (LED) [9] or computer screens (CS-SSVEP) [10]. These devices have achieved considerable success in BCI applications due to their stable refresh rates and reliable power supply. For instance, Chen et al. developed a 45-target SSVEP speller using an LED screen, achieving an accuracy of 84.1% and an information transfer rate (ITR) of 105 bits per minute with a stimulation duration of 2 seconds [11]. Wang et al. employed an LCD to develop a 40-command SSVEP paradigm for a BCI-based speller [12]. However, using CS-SSVEP requires users to frequently shift their focus between the monitor and the real-world visual environment, which can be highly inconvenient [13]. Therefore, CS-SSVEP exhibits certain limitations for practical applications.

Augmented reality-based SSVEP (AR-SSVEP) presents a promising alternative, enabling users to view stimuli within a real-world scenario through an AR device [14], [15]. This approach allows for more natural interaction with the real world via SSVEP, which could be the focal point for future research in SSVEP. However, it is hindered by lower accuracy and a limited number of target frequencies [16], [17]. Notably, the accuracy of AR-SSVEP declines as the number of target frequencies increases. With an increasing number of

frequencies, the smaller frequency interval is a major factor contributing to the decline of AR-SSVEP accuracy [17], [18]. On the other hand, the refresh rates of AR devices are not as stable as those of LED or computer screens [19], [20]. Such instability can prevent the target stimuli from achieving the expected frequencies, leading to misrecognition [21], [22]. The likelihood of incorrect frequency labeling increases when the frequency intervals are narrower. Therefore, increasing the number of targets by expanding the frequency range is not advisable for an AR setting. It is essential to address the limitations in the number of available targets while maintaining accuracy for AR-SSVEP.

Various methods have been proposed to increase the number of targets in SSVEP from different perspectives. Hwang et al. proposed a modified dual-frequency chessboard stimulation that combines four different frequencies to expand the number of targets to ten [23], achieving an accuracy of 87.23% with ten participants. However, the classification time windows ranged from 4 to 6 seconds, making this dual-frequency method inconvenient for practical applications due to the long stimulus duration required for accurate classification. Chen et al. introduced a spectrally-dense joint frequency-phase modulation (sJFPM) method, increasing the target number to 120 [24]. The average accuracy of sJFPM was 92.47% with a frequency interval of 0.1 Hz. Nonetheless, the prerequisite for achieving good results with sJFPM is ensuring that the stimulus can be presented steadily at the specified frequency, which is a challenging requirement for the AR environment. Some studies have utilized high-frequency stimuli to increase the target number, achieving relatively good results [25]. However, high-frequency SSVEP also requires a stable refresh rate of the device. On the other hand, it is feasible to expand SSVEP targets by introducing additional bio markers. Currently, the mainstream approach includes the use of event-related potential (ERP) components for auxiliary discrimination [26], [27]. Han et al. proposed a hybrid BCI that integrates SSVEP, P300, and motion visual evoked potential (mVEP), expanding the original 36 SSVEP targets to 216 [28]. However, ERP components are not stable enough without superimposed averaging. Soram Kim et al. employed AR-P300 for drone control [29], achieving satisfactory accuracy by averaging 20 epochs, which is time-consuming. Moreover, unlike in a laboratory environment, ERP is susceptible to being elicited by unpredictable stimuli in real-world scenarios [30], which can also impact discrimination. Therefore, for AR applications, it is necessary to introduce a more stable bio-markers. Overall, previous methods for target expansion are either time-consuming or impose high requirements on equipment, rendering them unsuitable for real-world AR environments. In addition, most research has primarily focused on increasing CS-SSVEP targets. An effective strategy for expanding AR-SSVEP targets remains to be developed.

Nevertheless, most SSVEP research primarily focuses on two-dimensional paradigms, such as plane flickers or chessboard. Meanwhile, three-dimensional SSVEP paradigms have gradually attracted more research attention in recent years [21], [31], [32]. Particularly, 3D stimuli possess additional properties such as direction, shape, and motion,unlike 2D stimuli which only have different frequencies. These multiple properties of 3D stimuli may elicit more distinct biomarkers under different conditions. Within the visual-evoked potential (VEP) research community, various vision-related EEG mechanisms such as covert attention [33], orientation perception of ambiguous 3D objects [34], and motion perception [35] have been clarified. Compared to simple frequency feature, the number of targets and performance of AR-SSVEP might be enhanced by incorporating other visually evoked biomarkers. For instance, Kelly et al. integrated spatial attention information into SSVEP and achieved significant improvements by combining SSVEP features with attention-related parieto-occipital alpha band modulations [36]. However, most SSVEP paradigms are still predominantly expanding targets by adding more frequencies. There are few studies on SSVEP that integrate it with other potential visual biomarkers induced by 3D objects.

Therefore, leveraging the advantages of 3D objects, a novel paradigm is proposed in this study. This paradigm incorporates 12 rotating cube flickers (6 frequencies and 2 rotation orientations), which may elicit both frequency and rotation-related bio-markers. To compare with the conventional method, a baseline experiment using a 2D paradigm consisting of 12 target frequencies was conducted. Our aim is to verify the feasibility of 3D paradigm in effectively enlarging the target number of AR-SSVEP. Compared to 2D paradigms, the 3D paradigm aims to achieve more accurate discrimination by integrating rotational and frequency information. In this study, 12 subjects were recruited in the experiment. Three methods, including filter bank canonical correlation analysis (FBCCA) [37], extended canonical correlation analysis (ECCA) [38], and task related component analysis (TRCA) [39] were employed for SSVEP classification. Specifically, a common spatial pattern (CSP) based squeeze-and-excitation convolutional neural networks (CSP-SECNN) [40] was innovatively applied to classify rotation-related visual EEG in this study. Finally, classification accuracy and ITR [41] were compared between 2D and 3D paradigm.

The contributions of this study are outlined as follows:

1) **Development of a novel 3D hybrid paradigm:** We proposed a novel 3D paradigm which merges SSVEP and EEG elicited by visual perception of different rotation pattern to increase the number of targets effectively.
2) **Building effective model for the classification of rotation-evoked EEG:** We innovatively leveraged CSP-SECNN for the classification of rotation-evoked EEG and got a good result.
3) **Demonstration of enhanced performance in AR environments:** Experimental results with 12 participants validate the better performance of the 3D paradigm than 2D paradigm in AR settings.

The rest of this paper is organized as follows: Section II introduces the experiment details. Section III describes the methodology employed for the classification of 2D and 3D paradigms. Section IV presents the classification results. Section V offers a discussion from different perspectives. Finally, Section VI concludes this study and outlines future work.

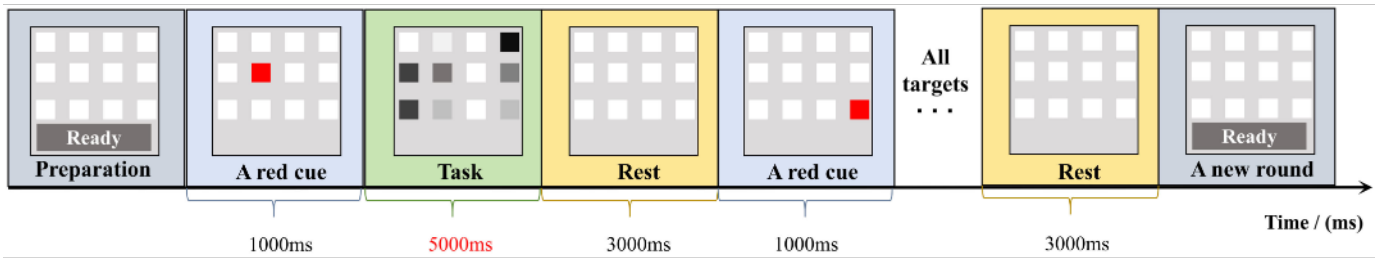

Fig. 1. Illustration for experiment procedure.

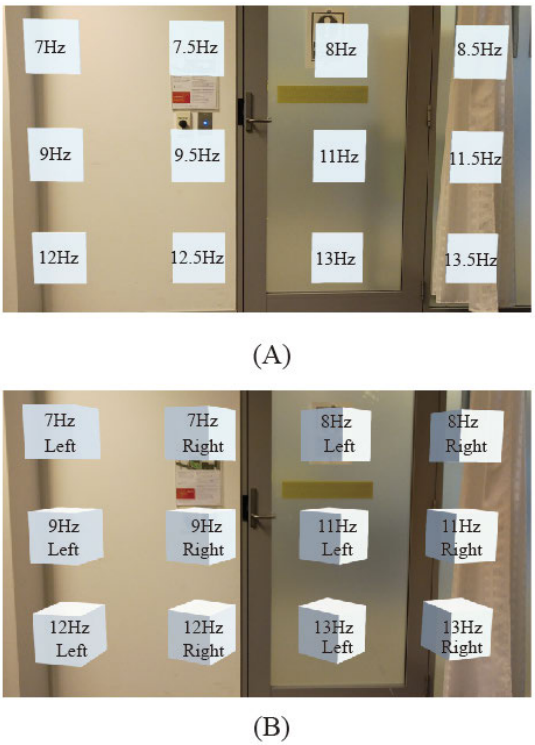

Fig. 2. Stimulation in Hololens-based AR environment: (A) Stimulation of 2D paradigm. (B) Stimulation of 3D paradigm.

II. Experiment

A. Stimulation Design

1) 3D Paradigm: For 3D stimulation (see 2(B)), 12 rotating cube flickers were selected as targets to elicit the corresponding frequency in EEG and patterns related to the visual perception of rotation. By combining six frequencies, including 7 Hz, 8 Hz, 9 Hz, 11 Hz, 12 Hz and 13 Hz with two rotation patterns (left and right rotation), twelve targets stimuli were generated with an approximate interval of 1 Hz. Ideally, the inclusion of a 10 Hz frequency would ensure a 1 Hz interval. However, the occipital alpha rhythm at 10 Hz could respond to other stimulation frequencies [42]. Specially, in response to external visual stimulation, EEG at 10 Hz could be elicited simultaneously with other target frequencies, potentially leading to misrecognition of SSVEP. Therefore, 10 Hz stimulus was excluded from this experiment.

The distance between the AR glasses and the stimulation was set at 3000 mm and the distance between adjacent flickers was 150 pixels horizontally and 135 pixels vertically. Each cube was sized at $472 \times 472 \times 472$ pixels, corresponding to a visual angle of 2.1°.The selection of these parameters is based on the experiment detailed in [14]. Sinusoidal coding was employed for the modulation of flicker frequency. The chroma value of the flickers was modulated by a sine wave at the target frequency. The color of the cube flickers alternated between black and white. The rotation speed of each flicker was set to 30 degree per second, as participants reported that this speed provided the most comfortable experience during pre-test.

2) 2D Paradigm: Twelve static flashing planes with different frequencies were selected as 2D stimulation (see 2 (A)), serving as a baseline paradigm for comparison with the 3D paradigm. The size of 2D stimuli is 472×472 pixels, which was same as 3D stimuli. The layout parameters of 2D paradigm were equal to the 3D one as well. As a baseline, the 2D paradigm expands the targets by increasing the number of frequencies. Accordingly, twelve frequencies including 7 Hz, 7.5 Hz, 8 Hz, 8.5 Hz, 9 Hz, 9.5 Hz, 11 Hz, 11.5 Hz, 12 Hz, 12.5 Hz, 13 Hz, 13.5 Hz were selected for 2D flickers, which had a smaller interval.

3) Paradigm Implement: Both 2D and 3D paradigm were implemented with Unity 3D (Unity Technologies, Inc) and presented with a HoloLens 2 headset (Microsoft Corporation). Initially, the paradigm projects were set up in Unity with the Universal Windows Platform (UWP). The Mixed Reality Toolkit (MRTK) was used for HoloLens features, and then these paradigms were built and deployed through Visual Studio to the HoloLens device. The refresh rate of HoloLens 2 is 60 Hz, which is high enough to support the stimuli display.

B. Experiment Procedure

The experimental procedure is depicted in Figure 1. Here, an illustration for 2D experiment is given to show the procedure details. The procedures for 2D and 3D experiment are completely same. The only difference between 2D and 3D experiment is the stimulation type. At the beginning of each round, a preparatory interface displaying twelve cubes and a 'Ready' virtual button was presented to the subjects. During this period, subjects were allowed to adjust their status and positions. This interface remained visible until the subject presses the 'Ready' button. After preparation, subject could start the experiment by pressing the 'Ready' button contactlessly with a finger gesture (Drag the AR cursor to the 'Ready' button with thump finger and index finger, then release these two fingers). Subsequently, one of the twelve cubes turned red to indicate the target. The cue period lasted 1000 ms, providing sufficient time for subjects to locate the target flicker and adjust their head posture. Following this, all twelve flickers flashed for 5000 ms and subjects were expected to fixate on the target without moving their eyes. This was

followed by a 3000 ms rest period. Each round comprised 12 trials, with each flicker serving as the target once, and ended once all flickers had been targeted. The sequence of target was random within a round. Subjects could start the next round at any time when they were ready.

To minimize the influence of participant fatigue and battery level fluctuations on the two experiments, the 3D and 2D paradigms were conducted alternately, switching every six rounds. Each participant completed 36 rounds in total, comprising 18 rounds each in 2D and 3D paradigms. Consequently, 216 samples were collected for each paradigm.

C. Participants and Data Acquisition

Twelve subjects, including 7 males and 5 females, participated in this experiment. They ranged in age from 22 to 32. Among them, three subjects had prior experience with SSVEP experiments, while the remainder participated for the first time. All subjects were mentally healthy and possessed either normal vision or corrected-to-normal vision. Prior to the experiment, written ethical consent, approved by the University of Technology Sydney's Ethical Committee (Grant number: UTS HREC REF No. ETH20-5371), was obtained from each participant. For data acquisition, EEG signals were collected using a 64-channel Neuroscan electroencephalograph cap, which followed the 10-20 electrode placement standard [43]. And a Synamps2 amplifier (Compumedics Neuroscan, Charlotte, NC, USA) was used to enhance the signal. The sampling rate was set at 1000 Hz. At the onset of each trial, a marker corresponding to the target was transmitted to CURRY 8 (the data acquisition software of Neuroscan) via UDP and recorded synchronously with the EEG data.

D. Data Preprocessing

Initially, the left and right mastoid electrodes (M1 and M2) were designated as the reference channels to re-reference the data from the other 62 channels. To eliminate the artifacts, such as electrooculography (EOG) and electromyogram (EMG), independent component analysis (ICA) [44] was employed to make unsupervised source separation with EEGLAB [45]. The EEG data were then decomposed into several independent components. With the 'ADJUST' toolbox, significant artifact components were identified and removed [46]. An infinite impulse response (IIR) band-pass filter ranging from 0.5 Hz to 90 Hz was employed for data filtering. And a 50 Hz notch filter was applied to eliminate power line interference. With the reference to event markers, EEG samples were extracted by slicing the EEG with a 5000 ms time window. Finally, 216 pre-processed samples, each with the dimension of 62×5000 were extracted for each paradigm.

III. Methods

In this study, an experiment was conducted to compare the conventional 2D paradigm with the proposed 3D paradigm. For the 2D SSVEP, three classification methods, including ECCA, TRCA and FBCCA, were employed. Regarding the 3D SSVEP, it is also necessary to develop a classification model for EEG signals related to the visual perception of different rotation orientations. By integrating the results from a 6-class SSVEP classifier and a 2-class rotation classifier, twelve targets were established, as shown in Supplementary Figure 1. The classification methods for the 2D and 3D paradigms are described as follows:

A. SSVEP Classifiers

1) FBCCA: FBCCA is an unsupervised method for SSVEP classification that utilizes the harmonic information and outperforms the original CCA. In FBCCA, EEG data are filtered into multiple frequency bands using a band-pass filter, followed by using CCA to calculate the correlation coefficients between the sub-band data and reference signals. Subsequently, these coefficients are weighted and summed. Finally, the target frequency is classified by voting with the weighted correlation coefficients. According to [47], the FBCCA approach primarily involves three hyperparameters: the number of sub-bands N, and a and b for calculating the weighting coefficients. In this study, the parameters for FBCCA were set as $N = 5$, $a = 1.25$ and $b = 0.25$.

2) ECCA: ECCA incorporates subject-specific reference signals into the CCA and has shown great performance in SSVEP classification. In ECCA, in addition to utilizing standard sinusoidal and cosinusoidal signals as references, the template obtained from averaging SSVEP signals across various frequencies for each subject are also employed as an individual-specific reference, thereby enhancing the recognition of individual SSVEP responses. The detail of ECCA can be found in [38]. Furthermore, in calculating ECCA, the filter-bank concept was also adopted, which involves calculating correlation coefficients for SSVEP signals across multiple frequency bands and weighting them to obtain the final coeffifient. The parameters for the filter bank are the same as those of FBCCA.

3) TRCA: TRCA extracts task-related components by maximizing the reproducibility during the task period. The parameters for the filter bank are same as that of FBCCA. In TRCA, a frequency-specific template is constructed by averaging the EEG data associated with each frequency. Spatial filters are subsequently derived through the TRCA method. The result is determined by calculating the correlation coefficients between the projected EEG data and multiple reference signals. The details of TRCA calculation can be found in [39].

Additionally, TRCA also incorporates a filter-bank mechanism to process EEG data across multiple frequency bands, with its parameters aligned consistently with those used in FBCCA.

B. Classifier for 3D Paradigm

Classification for 3D SSVEP is divided into two parts (see Supplementary Figure 1). EEG data are input simultaneously into the SSVEP classifier and the classifier for rotation-related EEG. The intersection ($\cap$) of the results from the two classifiers is used as the final output.

For SSVEP classifier, the previously mentioned three methods including ECCA, TRCA and FBCCA are also employed for 3D paradigm.

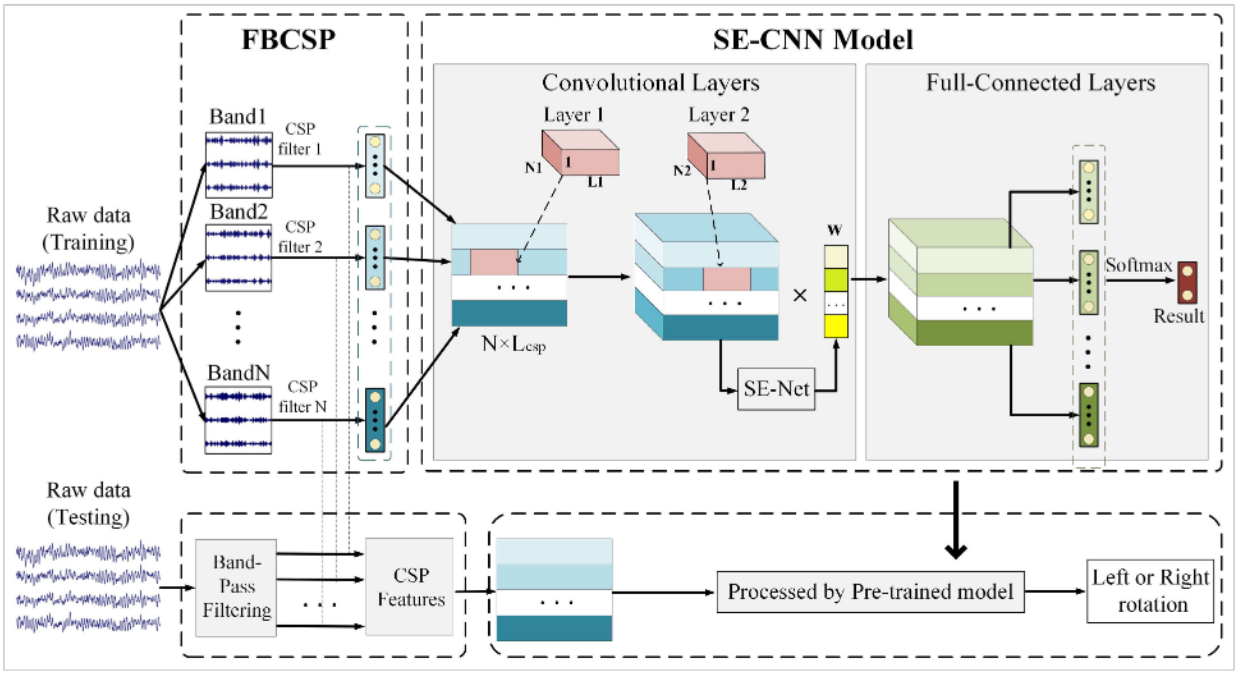

Fig. 3. Structure of CSP-SECNN model leveraged for rotation classification.

For rotation classification, CSP algorithm [48] was leveraged for feature extraction of EEG related to different rotations. The details of the feature extraction and discrimination processes for rotation are depicted in Figure 3.

CSP feature extraction is executed in two steps. Firstly, a set of spatial filters is derived. During this phase, covariance matrices for different classes of EEG data are computed. Subsequently, these covariance matrices are jointly diagonalized and whitened to obtain a mapping matrix, referred to as a CSP filter. The computation details of the CSP filter in this study are based on [48]. In the second step, EEG data are projected onto a common space with CSP filters and the variance difference of the projected data is enhanced. Thus, the variance of the projected data is selected as the feature for classification. The CSP feature f_n is calculated as Equation 1:

$$f_n = log(1 + var(WE_n)) \quad (1)$$

where E_n denote the EEG signal of the n-th class, W is the CSP filter. Based on previous work [49], CSP features derived from the first and last few rows of the CSP filters are distinctive. Typically, the number of the selected CSP filters ranges from 2 to 6. In this study, the first three and the last three rows of CSP filters were selected, as these six filters demonstrated optimal performance during test.

Next, the filter-bank idea was applied to feature extraction. Compared with CSP, Filter-bank Common Spatial Pattern (FBCSP) extracts features derived from multiple sub-bands, which leverages more useful information and can enhance classification performance [50]. In the FBCSP module, the original EEG data are initially filtered through N IIR bandpass filters and N sub-band signals are generated. Then N CSP filters are trained with these sub-band data to extract multi-band features.

Following feature extraction, SECNN is applied to discrimination. N features are stacked vertically and the input of CNN

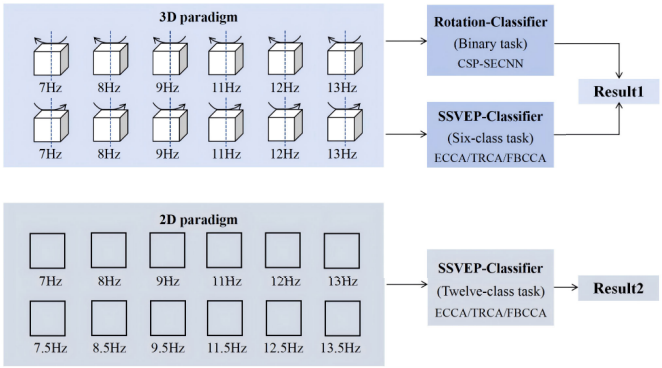

Fig. 4. Illustration of the comparison between 3D and 2D paradigm.

model is a $N \times L_{csp}$ feature map, where L_{csp} is the length of one CSP feature. Specifically, *Layer* 1 consists of two kernels where $N_1 = 1$ and $L_1 = 3$ while *Layer* 2 has four kernels where $N_2 = 2$ and $L_2 = 3$. Then, the SE-Net module is used for assigning a $1 \times N$ weight w adaptively to the feature of each band to enhance the useful features and suppress the redundant ones. The details of SE-Net can be found in [51]. Finally, the weighted features are combined with a Full-Connected layer, and a softmax layer is applied for the binary classification of rotation. In this study, Floating Point Operations (FLOPs) [52] was used for evaluating the computational complexity and calculated with the pytorch-flops-counter2.[1] The FLOPs of CSP-SECNN with 1 second data as input is 0.665 Mmac.

C. Performance Evaluation and Details

The purpose of this evaluation is to compare the performance of the proposed 3D paradigm with the conventional 2D paradigm (Figure 4). In the 3D paradigm, there are two

[1] https://github.com/sovrasov/flops-counter.pytorch

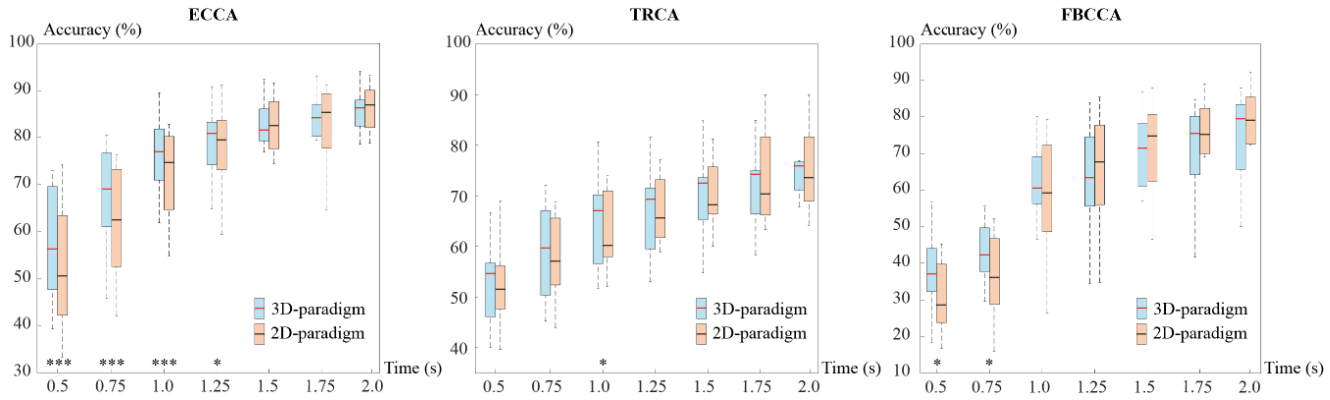

Fig. 5. Mean accuracy of 2D and 3D paradigm. In each subplot, the horizontal axis represents the length of the EEG signal in units of seconds.The vertical axis indicates accuracy. The blue bars indicate the accuracy of 12-class classification of 3D paradigm. From left to right, The results of the 3D paradigm were calculated with (ECCA, CSP-SECNN), (TRCA, CSP-SECNN), and (FBCCA, CSP-SECNN). The orange bars indicate the accuracy of 2D paradigm (12-class SSVEP). * indicates significant differences between accuracy of 3D and 2D paradigm, as assessed by paired t-test (*$p < 0.05$, ***$p < 0.001$).

parallel processing streams for EEG decoding. Data of 3D paradigm are input into both a binary rotation classifier and a six-class SSVEP classifier. The rotation classifier outputs the target orientation (left/right), and the SSVEP classifier outputs the target frequency. Taking the intersection of the results from the two classifiers as the final outcome, 12 targets can be generated. For the 2D paradigm, data are input into a SSVEP classifier, which outputs the classification result for 12 target frequencies.

According to [12], channels Pz, POz, PO3, PO4, PO5, PO6, O1, Oz and O2 are selected as the input channels for the SSVEP classifier. As few studies have clearly defined which brain regions and frequency bands are involved in the visual perception of rotation, several combinations of brain cortex areas (frontal, motor, parietal, occipital) and frequency bands (delta, theta, alpha, beta, and low gamma) were tested on rotation classifier. The results indicated that the optimal performance was achieved using the low gamma bands of the occipital and posterior parietal lobes. Consequently, EEG from 19 occipital-parietal channels, including P1-P8, Pz, PO3-PO8, POz, O1, Oz, and O2, are selected as the inputs for the rotation classifier. For feature extraction in the rotation classifier, six sub-bands ranging from 30 Hz to 45 Hz, segmented with an interval of 2.5 Hz, were utilized.

Except for FBCCA, the remaining methods are supervised learning models for which training and test datasets are necessary. Five-fold validation was implemented for evaluation,, where 80% of the data served as the training set and the remaining 20% as the test set. The length of the EEG segments used for testing was increased from 0.5s to 2.0s in increments of 0.25s.

In addition to accuracy, the ITR was also calculated for evaluation:

$$\mathrm{ITR} = \frac{60}{T_g + T_s + T_w} \left[\log_2 N + P\log_2 P + (1-P)\log_2 \frac{1-P}{N-1}\right] \quad (2)$$

where T_g is the gaze shifting time, T_s denotes the response latency, T_w denotes the duration of EEG input, N is the total number of targets and P is the accuracy. Here, the T_g was $0.5s$ and the T_s was $0.135s$ [37].

IV. Result

The average accuracies and ITR of 2D and 3D paradigm are presented in Figure 5 and Figure 6 respectively. With ECCA, it is evident that the 3D paradigm outperforms the 2D paradigm ($p < 0.05$) when the EEG segment length is shorter than 1.5s. With a 1s EEG segment, the 3D paradigm achieved an average accuracy of 76.50% and an ITR of 70.42 bits/min, while the accuracy and ITR of 2D paradigm were 72.07% and 62.28 bits/min respectively ($p_{acc} = 2.91 \times 10^{-4}$, $p_{ITR} = 8.13 \times 10^{-6}$, $a = 0.05$). With TRCA, the average accuracy and ITR of the 3D paradigm were higher than those of the 2D paradigm, although the difference was not statistically significant. The accuracies of both paradigms were less than 70% when the time length was 1.0s. Nevertheless, the 3D paradigm still outperforms the 2D one ($p_{acc} = 0.0287$, $a = 0.05$). Regarding FBCCA, both accuracy and ITR for the 3D paradigm were significantly higher than those for the 2D only when the EEG segment was less than 1s ($p_{acc0.5s} = 0.0175$, $p_{acc0.75s} = 0.0236$, $p_{ITR0.5s} = 0.0023$, $p_{ITR0.75s} = 0.0312$, $a = 0.05$)

By comparing the three SSVEP classification algorithms, it can be seen that ECCA had the best performance for both paradigms, achieving an average accuracy of 81.63% in 1.5s. When the time length was relatively short (less than 1s), ECCA and TRCA outperformed FBCCA ($p < 0.01, a = 0.05$). When the time window was long enough ($t > 1.25s$, ECCA significantly outperformed both TRCA and FBCCA ($p < 0.001$, $a = 0.05$), while the performance difference between the latter two methods was negligible ($p > 0.05$, $a = 0.05$). Compared to FBCCA, ECCA introduces individual SSVEP templates as supplementary reference signals. This enhances the similarity between individual SSVEP signals and the reference signals, improving the discriminability between target and non-target frequencies [53], thereby achieving better classification results. TRCA relies on phase differences, and it performs well only when there is a significant phase difference [54]. In this

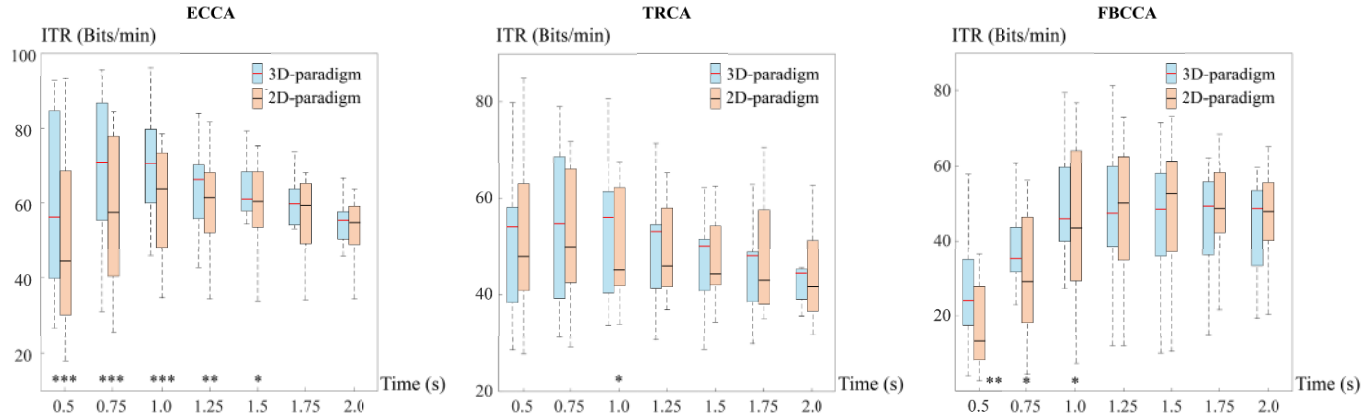

Fig. 6. Mean ITR of 2D and 3D paradigm. In each subplot, the horizontal axis represents the length of the EEG signal in units of seconds. The vertical axis indicates ITR in units of bits/min. The blue bars and orange bars indicate the ITR of 3D paradigm and 2D paradigm respectively. * indicates significant differences between ITR of 3D and 2D paradigm, as assessed by paired t-test (*$p < 0.05$, **$p < 0.01$,***$p < 0.001$).

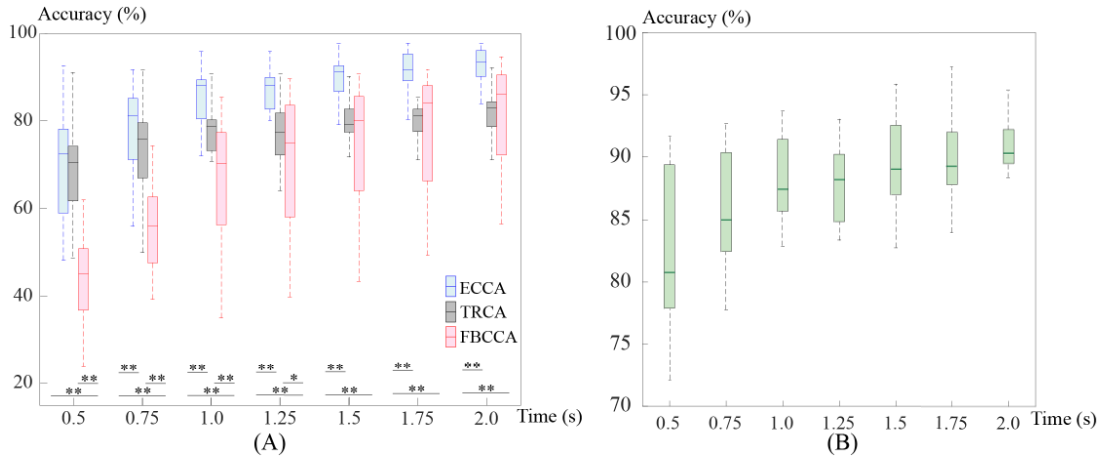

Fig. 7. Accuracy of two sub classifiers of 3D paradigm. In each subplot, the horizontal axis represents the length of the EEG signal in units of seconds.The vertical axis indicates accuracy. (A) Mean accuracy of three SSVEP classifiers (6-class SSVEP classification of 3D paradigm). * indicates significant differences between the accuracy of two methods, as assessed by paired t-test (*$p < 0.05$, **$p < 0.01$). (B) Mean accuracy of classifier of rotation (CSP-SECNN).

study, the phase among all stimuli was consistent, which may result in the poorer performance of TRCA. Overall, the mean accuracy of the best method (ECCA) illustrate that the 3D paradigm achieved better performance on shorter time segments (t<1.25) than the 2D paradigm. And when the time window is sufficiently long (t>1.25), there are only slight differences between the two approaches.

The accuracy of two sub-classifiers in the 3D paradigm is shown in Figure 7, which presents the classification results of SSVEP and rotation-induced EEG signals in the 3D paradigm respectively. ECCA significantly outperformed TRCA and FBCCA for SSVEP classification (($p < 0.01$, $a = 0.05$). The mean accuracy of TRCA was higher than that of FBCCA when the time length was less than 1.25s (($p < 0.01$, $a = 0.05$) and lower when the time length was longer ($t > 1.5s$).For rotation classification, the average accuracy reached 82.38% with a 0.5s time window and exceeded 90% when the time length was 2s. Moreover, the results of the 6-class 3D SSVEP classifier (Figure 7)(A) were consistent with those of 12-class classification of the 3D paradigm (Figure 5), where ECCA performed best. Thus, the result of sub-classifiers indicates that a good SSVEP classifier can improve the performance of 3D paradigm effectively.

Moreover, the individual results of both 3D and 2D paradigms are provided in supplementary Figure 2. S3 exhibited the best performance, achieving an accuracy of 80.56% and an ITR of 95.65 bit/min using the 3D paradigm with a duration of only 0.75 seconds. Individual results showed that not all subjects performed better in the 3D paradigm than in the 2D one. For S1 and S6, there is little difference in their performance between the 2D and 3D paradigms. Moreover, both subjects achieved relatively good results with these two paradigms. When the duration was 1s, the accuracy of S1 in the 2D and 3D paradigms was 80.42% and 81.86% respectively, while the accuracy of S6 reached 78.64% and 80.13% respectively. For S2, S4, S8, and S9, they achieved better accuracy and ITR with the 3D paradigm when the time window was shorter than 1.25s. And when time window was longer than 1.25s, their performance on the 3D paradigm was worse than the 2D paradigm. Take S2 as an example,

when time length was 0.75s, the accuracy was 76.85% for 3D paradigm and dropped to 70.31% for 2D paradigm. But with a time window of 1.25s, the accuracy of the 3D and 2D paradigms was 85.19% and 89.58%, respectively. For S3, S5, and S7, when the time window was shorter than 1.25s, their performance in the 3D paradigm was superior to that in the 2D paradigm. However, when the time window was longer, their performance was relatively similar in both paradigms. For S10, S11, and S12, regardless of the length of the EEG segment, they consistently achieved better performance on the 3D paradigm. Overall, except for S1 and S6, the remaining ten subjects could achieve better performance in the 3D paradigm with a short EEG segment. These findings are consistent to the group-level result shown in Figure 5 and 6, which imply that the performance of AR-SSVEP can be effectively improved in shorter time windows with the 3D paradigm.

V. Online Experiment

A. Experimental Details

To validate the feasibility of the proposed 3D paradigm and make a comparison to original 2D paradigm, this study conducted online experiments using the HoloLens 2 headset for stimulus presentation and a 64-channel Neuroscan EEG system for signal acquisition. The four participants who performed best in offline experiments participated in the online tests. Each participant's model was trained using data from their offline experiments. The sampling rate was set to 1000 Hz in online experiment. The selection of channels is consistent with the offline experiments. Considering the cross-session instability of EEG signals [55] and the interference of spontaneous EEG activity [56], the presentation time of stimuli in the online experiments was adjusted to 3 seconds to ensure stable elicitation of the corresponding EEG components.

Each experimental session included 12 trials, with each of the 12 stimuli randomly chosen as the target once, and participants were instructed to focus on the target stimuli during task period. After each stimulus presentation, EEG signals from the task phase were transmitted via TCP/IP [57] and preprocessed with a an IIR filter ranging from 0.5 Hz to 90 Hz before being input into classifiers. The results were recorded and compared with the labels to calculate the final accuracy. Each participant completed three rounds of online experiments, totaling 36 trials, with each stimulus being tested three times.

B. Online Result

The online result is shown in Table I. For 3D paradigm, three subjects achieved accuracies above 80%, with the highest result reaching 86.11%. The remaining participant whose accuracy was 77.78% also significantly exceeded the random baseline of 16.67%. Compared to the 2D paradigm, all subjects except subject 1 achieved better performance in the 3D paradigm. Especially for subject 4, his performance in the 2D paradigm was relatively poor. This result further indicates that as the target frequency increases, some participants with lower sensitivity to flicker frequency may face challenges.

TABLE I
The Result of Online Experiment for 4 Subjects

Subject ID	1	2	3	4	mean
3D Accuracy (%)	86.11	83.33	83.33	77.78	82.63
2D Accuracy (%)	88.89	77.78	80.56	66.67	78.48

The results of the online experiments to some extent validate the feasibility and outperformance of 3D paradigm. The discrepancy between online and offline results may stem from overfitting during offline model training and variations in EEG data across sessions. Therefore, addressing the cross-session transfer learning issues in the 3D paradigm can be a significant direction for future improvements.

VI. Discussion

In this section, some explanation work is given. Firstly, we analyze the rotation-evoked EEG patterns, providing explanations for why EEG signals induced by different rotations are distinguishable. Secondly, we discuss on two key factors which might affect the SSVEP classification: the introduction of rotation and the number of target frequencies. By exploring the impact of these factors, we explain why the 3D paradigm achieves superior results.

A. Pattern Analysis of Rotation-Evoked EEG

1) Spatial Filter Analysis: Firstly, we investigated the brain regions associated with rotation-evoked EEG. As mentioned above, CSP method was utilized for feature extraction in rotation classification. During this process, a set of spatial filters was derived, which could represent the active levels of the electrodes involved in classification [58]. Notably, motor imagery studies have shown that CSP filters assign high weights (absolute values) to task-dependent electrodes (e.g., C3 and C4 in the motor cortex) [58]. Some researchers selected electrodes of interest according to the weights [59]. Thus, by examining the distribution of CSP filters, it is possible to find the brain regions activated by the rotation task.

In detail, according to the CSP calculations [48], the number of CSP filters are equal to the number of electrodes. In this study, 62 electrodes were utilized, resulting in 62 CSP filters. Previous studies has indicated that the first and the last CSP filters can effectively represent the brain's activity levels under two distinct tasks [49], respectively. Therefore, the first and last CSP filters derived from rotation-evoked EEG (filtered between 30-45 Hz) are shown in Figure 8, illustrating the active brain areas during left and right rotations, respectively.

It is evident that the active electrodes are primarily located in the occipital and posterior parietal lobes, aligning with the regions typically associated with visual-evoked potentials. However, no significant differences were observed between the filter weights for left and right rotations (p>0.05, $a = 0.05$). Only minor contralateral differences were noted in the data for individual subject. For example, S11 and S12 exhibited high weights in the left posterior area during left rotation, and in the right posterior area during right rotation. Conversely, for subjects like S1 and S4, nearly all electrodes in the posterior brain regions showed high weights. Although no universal

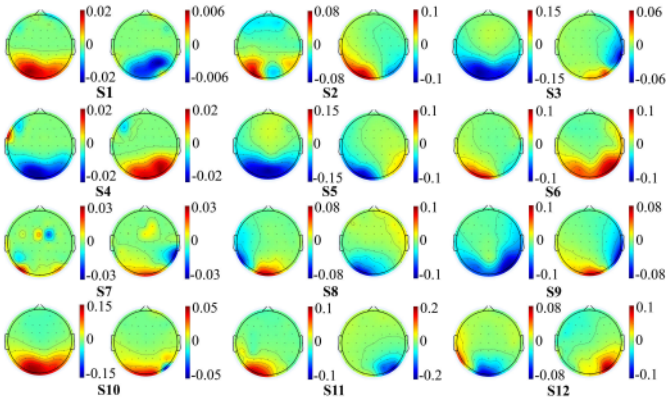

Fig. 8. Topological maps of CSP filters extracted from rotation classification. Left figure of each subject denotes the active area under left rotation. Right figure of each subject denotes the active area under right rotation. Dark blue and dark red denote higher negative and positive weights and both of them indicate highly active degree.

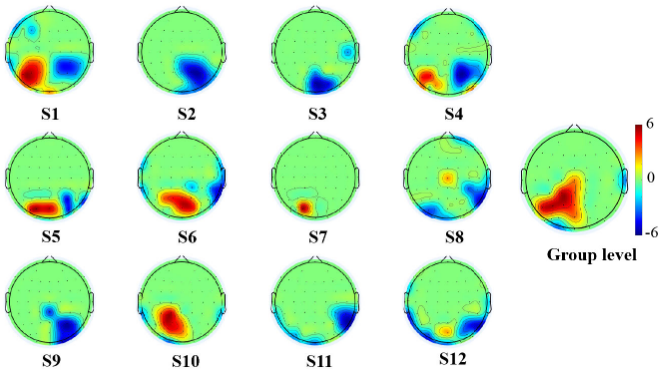

Fig. 9. Topological maps of significant PSD difference between left and right rotation. The values of all colorbars denote the t-value of comparison betweeen left and right rotation, ranging from -6 to 6. The group-level result was calculated with data of all subjects. The dark red and dark blue indicate high positive value and high negative value.

Algorithm 1 Pseudocode of Calculating PSD Difference Between EEG Elicited by Left Rotation (LR) and Right Rotation (RR) and Outputting Significant Results

Require: EEG_{LR}, EEG_{RR}

$PSD_{LR} \leftarrow Bandpower(EEG_{LR})$, $PSD_{RR} \leftarrow Bandpower(EEG_{RR})$

$p, t\_value_{L-R} \leftarrow ttest(PSD_{LR}, PSD_{RR})$

$p_{adj} \leftarrow FDR(p)$

$t\_value_{L-R} \leftarrow t\_value_{L-R} = 0 \; if \; p_{adj} > 0.05$

return t_value_{L-R}

distribution of filter weights has been established, these findings suggest that the occipital and posterior parietal lobes are relevant to rotation classification, offering guidance for electrode selection in practical applications of this paradigm.

2) Power Spectrum Density (PSD) Analysis: Considering that the visual perception of different rotations may elicit EEG synchronization or desynchronization like motor imagery, PSD analysis was conducted to elucidate the power difference between two rotation cases. In this study, the PSD difference for each electrode under left and right rotations was calculated, and the significant different value ($PSD_{left} - PSD_{right}$) were projected onto a topological map. Specifically, a data length of 2 seconds was used for PSD calculation. The details of calculation is shown in pseudocode 1, where $Bandpower$ is the function used to calculate the PSD of a specific frequency band, $ttest$ is the paired t-test that tests the significance of the differences between the PSD of EEG elicited by left and right rotation. The False Discovery Rate (FDR) was used to correct the results of t-test and prevent false positive results [60].

Individual results indicate significant PSD differences in the occipital and parietal areas. For subjects S1, S4, S5, S6, S7 and S10, the PSD in the left posterior cortex was significantly higher during left rotation than right rotation. Additionally, compared to left rotation, a significantly higher PSD in the right posterior regions was observed during right rotation in all subjects except for S7 and S10. Regarding the group level, significant PSD differences were only found in the left posterior brain region. Therefore, it can be concluded that when subjects viewed a cube rotating in left or right, EEG synchronization likely occurred in the ipsilateral posterior cortex. However, the PSD results across all subjects were not completely consistent due to inevitable individual differences, which is acceptable. Moreover, significant PSD differences were observed in the occipital and post-parietal cortex, consistent with the activated regions detected by CSP filters. This consistency provides an explanation for the effective classification of different rotation orientations.

Regarding this phenomenon, previous studies have suggested that the gamma response of the posterior brain may be associated with the visual cognition of moving objects [61], [62]. Müller et al. conducted an experiment where subjects were required to view a coherent stimulus (a single bar moving to the right) and an incoherent stimulus (two bars moving in opposite directions) [63]. They reported that gamma power increased significantly with the coherent stimulus compared to the incoherent one. Interestingly, higher gamma power was observed at the right parieto-occipital scalp sites (P4, O2, and T6), which aligned with our PSD results. However, they did not conduct experiments with bars moving to the left. A similar finding was reported by [62], where subjects viewed a grating moving horizontally from left to right. Their results showed a significantly higher gamma power in the right occipital-parietal lobes, consistent with our PSD analysis. In this study, the horizontal rotation might have a similar visual effect as the horizontal moving stimulation used in these studies and induce synchronization in unilateral brain areas. However, it is challenging to draw complete conclusions as few studies have reported gamma EEG patterns in response to moving visual stimuli in various directions. Nonetheless, we did observe significant contralateral differences in gamma power between the two rotation stimuli in our experiment, yielding good classification performance. Some studies suggest that gamma synchronization is related to bottom-up attentional processing [64], [65]. In this experiment, the behavior of watching a rotating flicker and generating associated EEG is also a form of bottom-up cognition. Rotating motion

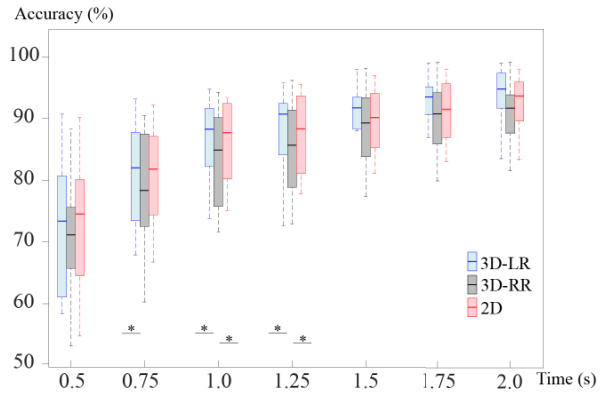

Fig. 10. Result of ablation evaluation on ECCA. The horizontal axis represents the length of the EEG signal in units of seconds.The vertical axis indicates accuracy. * indicates significant differences between the accuracy of two cases, as assessed by paired t-test ($*p < 0.05$).

may also induce a shift in the user's attention, potentially causing differences in gamma power across different posterior regions.

B. Analysis for the Superior Performance of 3D Paradigm

1) Effect of Introducing Rotation on SSVEP Classification:

An ablation evaluation was conducted to determine whether the introduction of rotation enhances or inhibits the SSVEP classification compared to the conventional stationary 2D flicker, which involved three classification tasks under three ablation conditions (see Supplementary Figure 3). Sub-dataset1 contained SSVEP elicited solely under left rotation (LR), and sub-dataset2 under right rotation (RR) only. Sub-dataset3 was separated from the 2D SSVEP dataset, which shares the same target frequencies as the 3D paradigm. All sub-datasets contained the same number of samples and target frequencies. According to 5 and 6, as ECCA had the best performance and demonstrated the most significant difference bewteen 2D and 3D paradigms, the ablation experiments utilized ECCA as the classification algorithm for SSVEP.

The result of ablation evaluation was presented in Figure 10, which indicated that the accuracy of 3D-RR on ECCA is lower than that of 3D-LR and 2D when the time segment is short ($t < 1.25s$). No significant difference was found between 2D and 3D-LR. Moreover, 2D SSVEP even performed significantly better than 3D-RR with time windows of 1s and 1.25s. Overall, compared to conventional 2D paradigm, there was no significant improvement for SSVEP classification of 3D paradigm. Thus, introducing rotation to flickers appears to have little effect on the discrimination of SSVEP, indicating that the superior performance of the 3D paradigm is not due to the involvement of rotation.

2) Effect of Increasing the Number of Target Frequencies:

To assess the impact of increasing the number of target frequencies on SSVEP decoding, confusion matrices for the 12-class 2D paradigm classification with ECCA are displayed in Figure 11. For 9 out of the 12 target frequencies, they were most likely to be misclassified as adjacent frequencies. For instance, 31 samples labeled as 12 Hz were misclassified as 11.5 Hz. Here, the fast Fourier transform (FFT) value

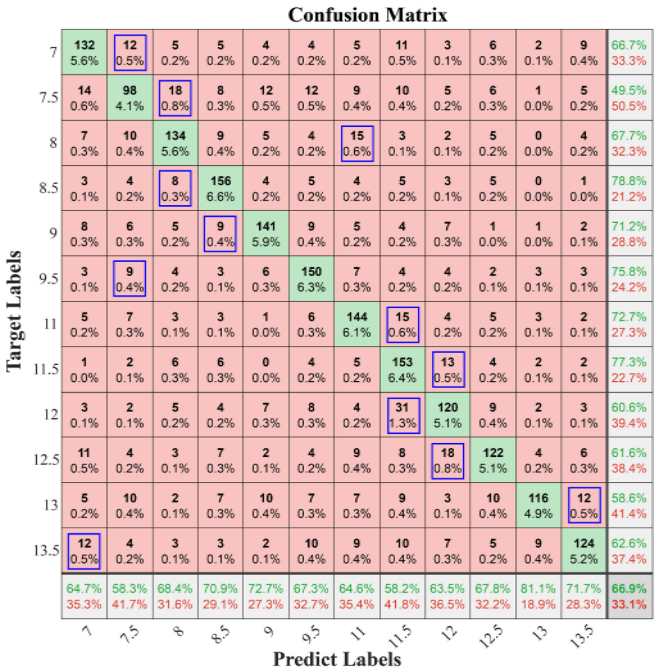

Fig. 11. Confusion matrix of 12-class 2D paradigm with 0.8s time length (summation of all 12 subjects). In the rightmost column of the figure, the green and red numbers respectively represent the precision and false discovery rate for each class. In the row at the bottom of the figure, the green and red numbers respectively represent the recall and false negative rate for each predicted class. The bold green and red numbers in the lower right corner of the figure represent the overall accuracy and error rate, respectively. The object enclosed by a blue box denotes the most misidentified frequency for one label.

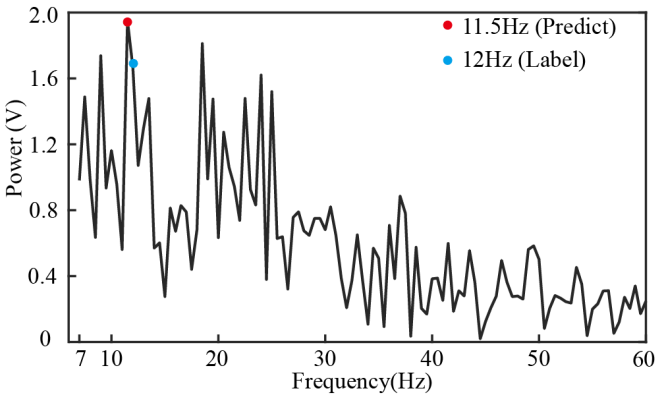

Fig. 12. The FFT results of error cases labeled as 12 Hz. The horizontal axis represents the frequency in units of Hz.The vertical axis indicates single-sided amplitude spectrum, with units in volts.

was computed for the Oz electrode data of all misclassified samples labeled as 12 Hz, and the average value was presented as Figure 12. It is evident that the frequency peak of the misclassified samples appears at 11.5 Hz rather than the expected 12 Hz. We suspect that this may be related to frame lost in the AR device, reducing the original 12 Hz stimulus to approximately 11.5 Hz. The misclassification of adjacent frequencies indicates that adding more target frequencies and reducing the interval between them may impair SSVEP classification. This finding partly explains why the accuracy of 2D paradigms significantly decreases when the number of

frequencies is increased from 6 to 12. In the 3D paradigm, fewer target frequencies were employed, reducing the adverse effects on classification caused by smaller frequency intervals. Meanwhile, the 6-class SSVEP and 2-class rotation classifiers maintained relatively high accuracy even with short EEG segments (as shown in Figure 7, when $t = 0.75s$, $ACC_{3D-SSVEP} = 81.58\%$, $ACC_{3D-rotation} = 85.75\%$). Consequently, by integrating these two sub-classifiers, the overall performance of the 12-class 3D paradigm surpassed that of the 2D paradigm.

The negative effects of increasing the frequency count can be explained by the unpredictability of device refresh rates [21], [22]. For instance, in the experiments conducted by Arpaia et al. [19], although the declared refresh rate of AR device was 30 Hz, the actual refresh rate fluctuated up to 32 Hz. This incorrect refresh rate caused the target frequency of 15 Hz to shift to 16 Hz, impacting the SSVEP accuracy. According to Figure 11, five target frequencies including 7 Hz, 7.5 Hz, 11 Hz, 11.5 Hz and 13 Hz were most frequently misclassified as their adjacent higher frequencies, which was consistent to [19]. Additionally, frame loss issues in Hololens were also observed in the AR-SSVEP experiments by Ke et al. [66]. Wang et al. further noted that using Hololens-based flickers is susceptible to refresh rate variability [67]. Moreover, their results demonstrated that AR-SSVEP was significantly less accurate than CS-SSVEP with shorter EEG segments ($t < 1.5s$). The result of the 12-class 2D paradigm in this study is also poor when EEG segments are short ($t < 1.25s$), which is consistent with their findings.

VII. Conclusion and Future Work

In this study, we proposed a 3D hybrid paradigm that combines visual rotation-evoked potentials and SSVEP to effectively expand the number of AR-SSVEP targets. When subjects viewed flickers rotating in different orientations, EEG components containing both frequency and rotation information were simultaneously induced, enabling the generation of more targets with their combination. We conducted a comparative experiment to assess the performance of this novel method against the conventional 2D paradigm that increased targets by adding more frequencies. ECCA was effectively applied to SSVEP classification, and a CSP-SECNN method was leveraged for rotation classification. Based on these two sub-classifiers, the proposed 3D paradigm achieved an ideal classification result. Twelve subjects participated in the experiment within an AR environment, and ten of them got enhanced performance with the 3D paradigm, particularly in shorter time windows. The mean accuracy and ITR further demonstrated the superior performance of the proposed 3D paradigm compared to the 2D approach.

We doubled the number of targets by introducing 2-class biomarkers elicited by the visual perception of rotation in this study. However, the current approach still has some limitations, which is only able to double the number of targets and can not expand more targets. In the future, we aim to incorporate additional biomarkers by investigating other properties of stimuli, such as various directions of rotation or different rotational speeds, to expand the targets in AR-SSVEP-based BCI systems more effectively.

References

[1] S. Samejima et al., "Brain-computer-spinal interface restores upper limb function after spinal cord injury," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1233–1242, 2021.

[2] X. Gao, D. Xu, M. Cheng, and S. Gao, "A BCI-based environmental controller for the motion-disabled," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, no. 2, pp. 137–140, Jun. 2003.

[3] M. Salvaris and F. Sepulveda, "Visual modifications on the P300 speller BCI paradigm," *J. Neural Eng.*, vol. 6, no. 4, Aug. 2009, Art. no. 046011.

[4] G. Townsend, B. Graimann, and G. Pfurtscheller, "Continuous EEG classification during motor imagery-simulation of an asynchronous BCI," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 12, no. 2, pp. 258–265, Jun. 2004.

[5] R. Kuś et al., "On the quantification of SSVEP frequency responses in human EEG in realistic BCI conditions," *PLoS ONE*, vol. 8, no. 10, Oct. 2013, Art. no. e77536.

[6] E. Yin, T. Zeyl, R. Saab, T. Chau, D. Hu, and Z. Zhou, "A hybrid brain–computer interface based on the fusion of P300 and SSVEP scores," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 23, no. 4, pp. 693–701, Jul. 2015.

[7] C. Farmaki, M. Krana, M. Pediaditis, E. Spanakis, and V. Sakkalis, "Single-channel SSVEP-based BCI for robotic car navigation in real world conditions," in *Proc. IEEE 19th Int. Conf. Bioinf. Bioeng. (BIBE)*, Oct. 2019, pp. 638–643.

[8] S. Park, H.-S. Cha, and C.-H. Im, "Development of an online home appliance control system using augmented reality and an SSVEP-based brain–computer interface," *IEEE Access*, vol. 7, pp. 163604–163614, 2019.

[9] C. Liu, M. Duan, Z. Duan, A. Liu, Z. Lu, and H. Wang, "An SSVEP-based BCI with LEDs visual stimuli using dynamic window CCA algorithm," *Biomed. Signal Process. Control*, vol. 76, Jul. 2022, Art. no. 103727.

[10] A. Luo and T. J. Sullivan, "A user-friendly SSVEP-based brain–computer interface using a time-domain classifier," *J. Neural Eng.*, vol. 7, no. 2, Apr. 2010, Art. no. 026010.

[11] X. Chen, Z. Chen, S. Gao, and X. Gao, "A high-ITR SSVEP-based BCI speller," *Brain-Comput. Interface*, vol. 1, nos. 3–4, pp. 181–191, Oct. 2014.

[12] 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, Oct. 2017.

[13] R. Zhang et al., "Improving AR-SSVEP recognition accuracy under high ambient brightness through iterative learning," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 1796–1806, 2023.

[14] X. Zhao, C. Liu, Z. Xu, L. Zhang, and R. Zhang, "SSVEP stimulus layout effect on accuracy of brain-computer interfaces in augmented reality glasses," *IEEE Access*, vol. 8, pp. 5990–5998, 2020.

[15] H. Si-Mohammed et al., "Towards BCI-based interfaces for augmented reality: Feasibility, design and evaluation," *IEEE Trans. Vis. Comput. Graphics*, vol. 26, no. 3, pp. 1608–1621, Mar. 2020.

[16] 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.

[17] R. Zhang et al., "The effect of stimulus number on the recognition accuracy and information transfer rate of SSVEP–BCI in augmented reality," *J. Neural Eng.*, vol. 19, no. 3, Jun. 2022, Art. no. 036010.

[18] D. Zhu, J. Bieger, G. Garcia Molina, and R. M. Aarts, "A survey of stimulation methods used in SSVEP-based BCIs," *Comput. Intell. Neurosci.*, vol. 2010, no. 1, pp. 1–12, 2010.

[19] P. Arpaia, E. De Benedetto, L. De Paolis, G. D'Errico, N. Donato, and L. Duraccio, "Performance enhancement of wearable instrumentation for AR-based SSVEP BCI," *Measurement*, vol. 196, Jun. 2022, Art. no. 111188.

[20] L. Angrisani, P. Arpaia, E. De Benedetto, L. Duraccio, F. Lo Regio, and A. Tedesco, "Wearable brain-computer interfaces based on steady-state visually evoked potentials and augmented reality: A review," *IEEE Sensors J.*, vol. 23, no. 15, pp. 16501–16514, 2023.

[21] H. Liu et al., "A comparative study of stereo-dependent SSVEP targets and their impact on VR-BCI performance," *Frontiers Neurosci.*, vol. 18, Apr. 2024, Art. no. 1367932.

[22] Y. Mustafa, M. Elmahallawy, T. Luo, and S. Eldawlatly, "A brain-computer interface augmented reality framework with auto-adaptive SSVEP recognition," in *Proc. IEEE Int. Conf. Metrology eXtended Reality, Artif. Intell. Neural Eng. (MetroXRAINE)*, Oct. 2023, pp. 799–804.

[23] H.-J. Hwang, D. Hwan Kim, C.-H. Han, and C.-H. Im, "A new dual-frequency stimulation method to increase the number of visual stimuli for multi-class SSVEP-based brain–computer interface (BCI)," *Brain Res.*, vol. 1515, pp. 66–77, Jun. 2013.

[24] X. Chen, B. Liu, Y. Wang, and X. Gao, "A spectrally-dense encoding method for designing a high-speed SSVEP-BCI with 120 stimuli," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 2764–2772, 2022.

[25] A. Chabuda, P. Durka, and J. Zygierewicz, "High frequency SSVEP-BCI with hardware stimuli control and phase-synchronized comb filter," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 2, pp. 344–352, Feb. 2018.

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

[27] M. Xu, H. Qi, B. Wan, T. Yin, Z. Liu, and D. Ming, "A hybrid BCI speller paradigm combining P300 potential and the SSVEP blocking feature," *J. Neural Eng.*, vol. 10, no. 2, Apr. 2013, Art. no. 026001.

[28] J. Han, M. Xu, X. Xiao, W. Yi, T.-P. Jung, and D. Ming, "A high-speed hybrid brain-computer interface with more than 200 targets," *J. Neural Eng.*, vol. 20, no. 1, Feb. 2023, Art. no. 016025.

[29] S. Kim, S. Lee, H. Kang, S. Kim, and M. Ahn, "P300 brain–computer interface-based drone control in virtual and augmented reality," *Sensors*, vol. 21, no. 17, p. 5765, Aug. 2021.

[30] P. Johnston et al., "Temporal and spatial localization of prediction-error signals in the visual brain," *Biol. Psychol.*, vol. 125, pp. 45–57, Apr. 2017.

[31] L. Niu et al., "Effect of 3D paradigm synchronous motion for SSVEP-based hybrid BCI-VR system," *Med. Biol. Eng. Comput.*, vol. 61, no. 9, pp. 2481–2495, Sep. 2023.

[32] S. R. Zehra, J. Mu, B. V. Syiem, A. N. Burkitt, and D. B. Grayden, "Evaluation of optimal stimuli for SSVEP-based augmented reality brain-computer interfaces," *IEEE Access*, vol. 11, pp. 87305–87315, 2023.

[33] L. Tonin, R. Leeb, A. Sobolewski, and J. D. R. Millán, "An online EEG BCI based on covert visuospatial attention in absence of exogenous stimulation," *J. Neural Eng.*, vol. 10, no. 5, Oct. 2013, Art. no. 056007.

[34] J. Kornmeier, M. Pfaffle, and M. Bach, "Necker cube: Stimulus-related (low-level) and percept-related (high-level) EEG signatures early in occipital cortex," *J. Vis.*, vol. 11, no. 9, p. 12, Aug. 2011.

[35] Y. Jiang et al., "Binding 3-D object perception in the human visual cortex," *J. Cognit. Neurosci.*, vol. 20, no. 4, pp. 553–562, Apr. 2008.

[36] S. P. Kelly, E. C. Lalor, R. B. Reilly, and J. J. Foxe, "Visual spatial attention tracking using high-density SSVEP data for independent brain-computer communication," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 13, no. 2, pp. 172–178, Jun. 2005.

[37] 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.

[38] C. M. Wong et al., "Learning across multi-stimulus enhances target recognition methods in SSVEP-based BCIs," *J. Neural Eng.*, vol. 17, no. 1, Jan. 2020, Art. no. 016026.

[39] 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.

[40] B. Cao, H. Niu, J. Hao, and G. Wang, "Building EEG-based CAD object selection intention discrimination model using convolutional neural network (CNN)," *Adv. Eng. Informat.*, vol. 52, Apr. 2022, Art. no. 101548.

[41] B. Obermaier, C. Neuper, C. Guger, and G. Pfurtscheller, "Information transfer rate in a five-classes brain-computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 9, no. 3, pp. 283–288, Sep. 2001.

[42] C. S. Herrmann, "Human EEG responses to 1–100 Hz flicker: Resonance phenomena in visual cortex and their potential correlation to cognitive phenomena," *Exp. Brain Res.*, vol. 137, nos. 3–4, pp. 346–353, Apr. 2001.

[43] R. Oostenveld and P. Praamstra, "The five percent electrode system for high-resolution EEG and ERP measurements," *Clin. Neurophysiol.*, vol. 112, no. 4, pp. 713–719, Apr. 2001.

[44] T.-W. Lee and T.-W. Lee, *Independent Component Analysis*. Cham, Switzerland: Springer, 1998.

[45] A. Delorme and S. Makeig, "EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *J. Neurosci. Methods*, vol. 134, no. 1, pp. 9–21, Mar. 2004.

[46] A. Mognon, J. Jovicich, L. Bruzzone, and M. Buiatti, "ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features," *Psychophysiology*, vol. 48, no. 2, pp. 229–240, Feb. 2011.

[47] T. Lee, S. Nam, and D. J. Hyun, "Adaptive window method based on FBCCA for optimal SSVEP recognition," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 78–86, 2023.

[48] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 8, no. 4, pp. 441–446, Dec. 2000.

[49] F. Jamaloo and M. Mikaeili, "Discriminative common spatial pattern sub-bands weighting based on distinction sensitive learning vector quantization method in motor imagery based brain-computer interface," *J. Med. Signals Sensors*, vol. 5, no. 3, p. 156, 2015.

[50] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, "Filter bank common spatial pattern (FBCSP) in brain-computer interface," in *Proc. IEEE Int. Joint Conf. Neural Netw. (IEEE World Congr. Comput. Intell.)*, Jun. 2008, pp. 2390–2397.

[51] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141.

[52] M. E. Paoletti, J. M. Haut, X. Tao, J. Plaza, and A. Plaza, "FLOP-reduction through memory allocations within CNN for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 7, pp. 5938–5952, Jul. 2021.

[53] M. Nakanishi, Y. Wang, Y.-T. Wang, and T.-P. Jung, "A comparison study of canonical correlation analysis based methods for detecting steady-state visual evoked potentials," *PLoS ONE*, vol. 10, no. 10, Oct. 2015, Art. no. e0140703.

[54] M. Nakanishi, Y.-T. Wang, and T.-P. Jung, "Optimizing phase intervals for phase-coded SSVEP-based BCIs with template-based algorithm," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2018, pp. 650–655.

[55] Y. Zhu, Y. Li, J. Lu, and P. Li, "EEGNet with ensemble learning to improve the cross-session classification of SSVEP based BCI from ear-EEG," *IEEE Access*, vol. 9, pp. 15295–15303, 2021.

[56] 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.

[57] L. Parziale et al., *TCP/IP Tutorial and Technical Overview*. NY, USA: IBM Redbooks, 2006.

[58] A. P. Costa, J. S. Møller, H. K. Iversen, and S. Puthusserypady, "An adaptive CSP filter to investigate user independence in a 3-class MI-BCI paradigm," *Comput. Biol. Med.*, vol. 103, pp. 24–33, Dec. 2018.

[59] N. Masood, H. Farooq, and I. Mustafa, "Selection of EEG channels based on spatial filter weights," in *Proc. Int. Conf. Commun., Comput. Digit. Syst. (C-CODE)*, Mar. 2017, pp. 341–345.

[60] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: A practical and powerful approach to multiple testing," *J. Roy. Stat. Soc., Ser. B*, vol. 57, no. 1, pp. 289–300, Jan. 1995.

[61] E. V. Orekhova, A. V. Butorina, O. V. Sysoeva, A. O. Prokofyev, A. Y. Nikolaeva, and T. A. Stroganova, "Frequency of gamma oscillations in humans is modulated by velocity of visual motion," *J. Neurophysiol.*, vol. 114, no. 1, pp. 244–255, Jul. 2015.

[62] S. D. Muthukumaraswamy and K. D. Singh, "Visual gamma oscillations: The effects of stimulus type, visual field coverage and stimulus motion on MEG and EEG recordings," *NeuroImage*, vol. 69, pp. 223–230, Apr. 2013.

[63] M. Müller et al., "Visually induced gamma-band responses in human electroencephalographic activity ? A link to animal studies," *Exp. Brain Res.*, vol. 112, no. 1, pp. 96–102, Nov. 1996.

[64] W. Lutzenberger, F. Pulvermüller, T. Elbert, and N. Birbaumer, "Visual stimulation alters local 40-Hz responses in humans: An EEG-study," *Neurosci. Lett.*, vol. 183, nos. 1–2, pp. 39–42, Jan. 1995.

[65] M. M. Müller and T. Gruber, "Induced gamma-band responses in the human EEG are related to attentional information processing," *Vis. Cognition*, vol. 8, nos. 3–5, pp. 579–592, Jun. 2001.

[66] 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.

[67] Y. Wang, K. Li, X. Zhang, J. Wang, and R. Wei, "Research on the application of augmented reality in SSVEP-BCI," in *Proc. 6th Int. Conf. Comput. Artif. Intell.*, Apr. 2020, pp. 505–509.