

# WAVEFORM-BASED MULTI-STIMULUS CODING FOR BRAIN-COMPUTER INTERFACES BASED ON STEADY-STATE VISUAL EVOKED POTENTIALS

Yutaro Tanji<sup>†</sup>, Masaki Nakanishi<sup>‡</sup>, Kaori Suefusa<sup>†</sup>, and Toshihisa Tanaka<sup>†,§</sup>

<sup>†</sup> Tokyo University of Agriculture and Technology, Tokyo, Japan

<sup>‡</sup> University of California San Diego, La Jolla, CA, USA

<sup>§</sup> RIKEN Brain Science Institute, Saitama, Japan

## ABSTRACT

Multiple stimulus coding plays an important role in a steady-state visual evoked potential (SSVEP)-based brain-computer interface (BCI). In conventional SSVEP-based BCIs, multiple visual stimuli are modulated with different properties such as frequencies and/or phases. However, the number of properties that can be assigned to visual stimuli rendered on a computer monitor is always limited by its refresh rate, leading to a system with limited commands or functions. To alleviate this issue, this study proposes a novel waveform-based stimulus coding method that uses modulation waveforms to differentiate resulting SSVEPs. In this paper, the discriminability of 12-class SSVEPs modulated by three types of waveforms (i.e., rectangle, sinusoidal and triangle waveforms) and four frequencies (i.e., 12, 13, 14, and 15 Hz) was investigated by computing its classification accuracy. The results showed the SSVEPs modulated by different waveforms can be successfully distinguished when using the state-of-the-art canonical correlation analysis (CCA)-based method with an average accuracy of 92.31%. This result suggests that the proposed method has great potential to significantly increase the number of functions in an SSVEP-based BCI system.

**Index Terms**— Brain-computer interfacing (BCI), Electroencephalography (EEG), Steady-state visual evoked potentials (SSVEP)

## 1. INTRODUCTION

Steady-state visual evoked potentials (SSVEPs) are brain's electrical oscillations elicited by repetitive visual stimulation, which can be measured by using electroencephalography (EEG) [1]. An SSVEP is known as a photic driving response characterized by a sinusoidal-like waveform at the stimulus frequency and its harmonics. Due to its robust characteristics, SSVEPs have been widely used in neuroengineering [1] and visual neuroscience [2]. In particular, a brain-computer interface (BCI), which provides a direct communication pathway

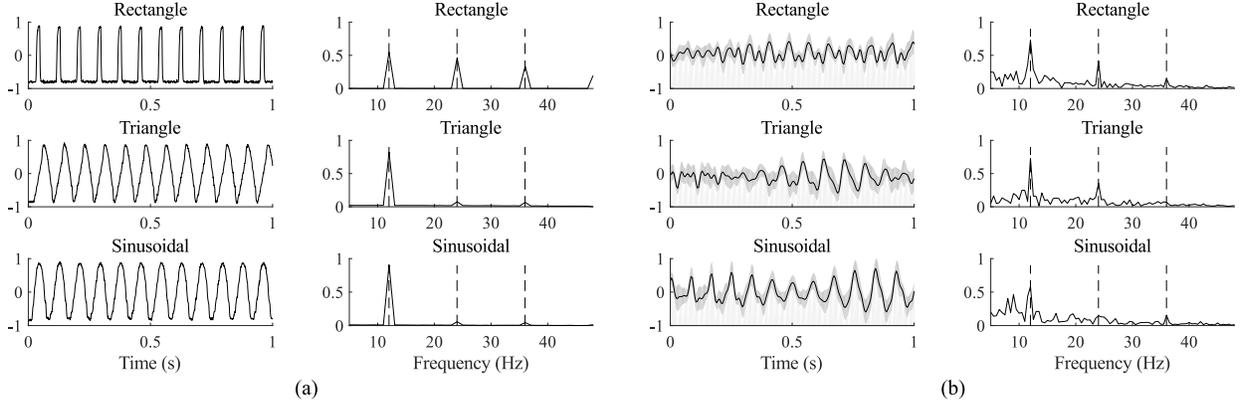
between human brain and external devices, has been the most popular application of SSVEPs [3].

In a traditional SSVEP-based BCI, users gaze at one of multiple visual stimuli modulated by different stimulus frequencies, resulting in SSVEPs that exhibit the same frequency as the target stimulus [4]. The target stimulus, which the user is gazing at, can be identified through analyzing the recorded SSVEPs. In this way, the system can indirectly translate users' intentions into commands for controlling external devices. Since the first idea of SSVEP-based BCI was depicted in 1970s [5], many researchers have attempted to improve its performance by addressing the following issues: 1) multi-target coding method, and 2) target identification algorithm [6]. A series of researches has shown a remarkable improvement in the performance of SSVEP-based BCIs in the past decade [7].

It has been a main challenge in a practical SSVEP-based BCI to increase the number of visual stimuli without compromising the discriminability (i.e., classification accuracy) of elicited SSVEPs. In general, stimulus frequencies need to be selected within a narrow frequency range since the signal-to-noise ratio (SNR) of SSVEPs are inconsistent across different frequency ranges [8]. Therefore, to realize a large number of BCI commands, a high frequency resolution (i.e., a small interval between two adjacent stimulus frequencies) is required. However, it is well known that increasing frequency resolution negatively affect in its classification accuracy [9, 10]. For example, in two separate studies, the systems with different frequency resolutions showed significantly different accuracies (Interval: Accuracy; 1.0 Hz: > 90% [9]; 0.2 Hz: < 30% [10]) even with the identical target identification approach. This problem could be solved by combining additional modulation patterns with frequencies to modulate visual stimuli. For example, the frequency shift keying (FSK) [11, 12] have been succeeded in increasing the number of BCI commands using a small number of frequencies. In several other studies, the efficacy of hybrid frequency and phase coding methods has also been demonstrated [13, 6, 7].

This study proposes a novel stimulus modulation method in which visual stimuli are modulated by different waveforms

This research was supported in part by JSPS KAKENHI Grant Number 15H04002. Email: tanakat@cc.tuat.ac.jp



**Fig. 1.** (a) Stimulus signals recorded by a photodiode and its amplitude spectra based on rectangle, triangle, and sinusoidal waves at 12 Hz, and (b) time series and amplitude spectra of elicited SSVEPs after the CCA-based spatial filtering corresponding to each stimulus waveform recorded from a subject (s3). The dashed lines in the spectra indicate the fundamental and harmonics frequencies. The time series SSVEPs obtained by averaging all trials after narrow band-pass filtering.

including rectangle, sinusoidal, and triangle waves. Since those waveforms consist of different combination of fundamental and harmonic frequency components [14], resulting SSVEPs might have consistent components to the stimulus waveforms with sufficient discriminability to each other. Teng et al. have compared the frequency classification accuracy using different stimulation waveforms and revealed that the rectangle wave with 50% duty cycle achieved the highest accuracy [14]. However, to our knowledge, the feasibility of stimulus modulation using waveforms as visual target for BCI has never been explored. In this paper, the classification accuracy of SSVEPs tagged with different waveforms and frequencies was evaluated using two target identification algorithms. The feasibility of applying the proposed method to an online system was also discussed by estimating an information transfer rate (ITR).

## 2. MATERIALS AND METHODS

### 2.1. Stimulus Modulation

This study employs different waveforms including rectangle, sinusoidal, and triangle waves to modulate multiple visual stimuli. The luminance changes of the stimuli can be modulated by stimulus sequences whose dynamic range is  $[0, 1]$ , where 0 represents black and 1 represents white. The stimulus sequences based on rectangle  $s_{\text{rect}}(f, i)$ , sinusoidal  $s_{\text{sin}}(f, i)$  and triangle  $s_{\text{tri}}(f, i)$  waveforms at stimulus frequency  $f$  can be generated by the following equations:

$$s_{\text{rect}}(f, i) = \text{square} [2\pi f (i/f_r)] \quad (1)$$

$$s_{\text{sin}}(f, i) = \frac{1}{2} \{1 + \sin [2\pi f (i/f_r)]\} \quad (2)$$

$$s_{\text{tri}}(f, i) = \text{triangle} [2\pi f (i/f_r)] \quad (3)$$

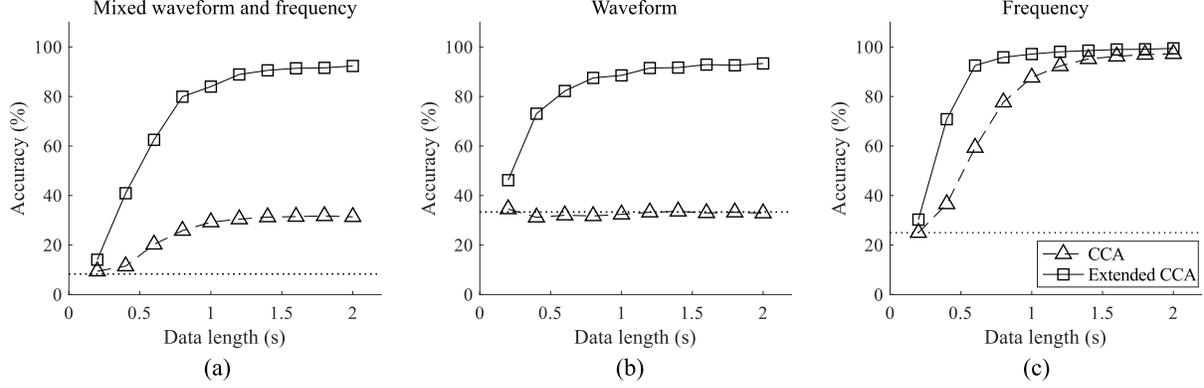
where,  $i$  indicates the frame index, and  $f_r$  indicates the refresh rate of a monitor. The duty cycle of the rectangle waves was set to 25% in this study. Fig. 1(a) shows the stimulus signals recorded by a photodiode placed over visual stimuli modulated by the three waveforms at 12 Hz presented on a computer monitor and its amplitude spectra.

### 2.2. Data Acquisition

A ViewPixx 3D 23-inch liquid crystal display (LCD) screen (VPixx Technologies, Inc.) with a refresh rate of 120 Hz was used to present visual stimuli. Four visual flickers (each with a size of 5.23 cm  $\times$  5.23 cm) with different frequencies from 12 to 15 Hz with an interval of 1 Hz were presented on the screen. The flickers were horizontally aligned with an interval between two neighboring stimuli of 5.23 cm. The modulation waveforms were the same across four stimuli, but they can be adjustable for each experimental session. The stimulation program was developed under MATLAB (Mathworks, Inc.) using the Psychophysics Toolbox Extensions [15].

Six healthy male adults (mean age: 22.8 years) with normal or corrected-to-normal vision participated in this study. Before participating in the experiment, all the subjects were asked to read and sign a written informed consent form approved by the research ethics committee of Tokyo University of Agriculture and Technology. In the experiment, the subjects were instructed to gaze at one of the stimuli indicated by the stimulus program for 3 s followed by a 3-s short break. In each block, the subjects completed four trials corresponding to all four stimuli. The modulation waveform was selected randomly for each block, and was fixed through trials in a block. Each subject performed 15 blocks for each waveform.

EEG data were acquired using g.SCRABEO Ag/AgCl active electrodes (g.tec medical engineering GmbH) placed over occipital area with reference to A1 and ground at AFz. EEG



**Fig. 2.** Averaged accuracy of (a) mixed waveform and frequency, (b) waveform, and (c) frequency classification across subjects as a function of data length from 0.2 to 2.0 s. The dotted lines indicate chance-level accuracy in classification ((a) 8.33 %, (b) 33.33 %, and (c) 25.00 %).

signals were amplified by MEG-6116 (Nihon Kohden, Corp.) and digitized by AIO-163202FX-USB (Contec, Co.) at a sampling rate of 1,200 Hz. Recorded EEG data were band-pass filtered between 10–50 Hz, and then 3-s data epochs corresponding to the stimulus duration were extracted. In the following analysis, the EEG signals recorded at PO8, O1, Oz, O2, and Iz were used to calculate the classification performance.

### 2.3. Target Identification Algorithms

#### 2.3.1. CCA-Based Method

Canonical correlation analysis (CCA), which is a statistical method to measure the underlying correlation between two sets of multidimensional variables, has been widely used to detect the frequency of SSVEPs [16, 17]. Considering two multidimensional variables  $S_1$  and  $S_2$ , CCA finds the weight vectors  $w_1$  and  $w_2$  which maximize the correlation between the projected signals  $w_1^T S_1$  and  $w_2^T S_2$  by solving the following problem:

$$\rho = \max_{w_1, w_2} \frac{E[w_1^T S_1 S_2^T w_2]}{\sqrt{E[w_1^T S_1 S_1^T w_1] E[w_2^T S_2 S_2^T w_2]}}. \quad (4)$$

The maximum of  $\rho$  with respect to  $w_1$  and  $w_2$  is the maximum canonical correlation. In the detection of SSVEPs, reference signals  $Y_{f,l} \in \mathbb{R}^{N_s \times N_h}$  consist of sets of stimulus sequences first need to be generated as follows:

$$Y_{f,l} = [s_l(f, i), s_l(2f, i), \dots, s_l(N_h f, i)] \quad (5)$$

where,  $l \in \{\text{rect}, \text{sin}, \text{tri}\}$  indicates the type of modulation waveforms,  $N_s$  is the number of sample points, and  $N_h$  is the number of harmonics ( $N_h = 3$  in this study). CCA calculates the canonical correlation  $\rho_{f,l}$  between the  $N_c$ -channel EEG signals  $X \in \mathbb{R}^{N_c \times N_c}$  and reference signals  $Y_{f,l}$  at each

stimulus frequency and waveform. The pair of target stimulus frequency and waveform ( $f^*, l^*$ ) can be identified by the following equation:

$$f^*, l^* = \arg \max_{f,l} \rho_{f,l} \quad (6)$$

#### 2.3.2. Extended CCA-Based Method

The extended CCA-based method that incorporates individual calibration data has shown improved performance in detecting frequency and phase of SSVEPs compared with the CCA-based method [6]. The method finds three spatial filters which enhance the SSVEP components. The three spatial filters are obtained as weight vectors by applying CCA to all two combinations of three multichannel signals: 1) a test EEG signal  $X$ , 2) reference signals  $Y_{f,l}$  in (5), and 3) individual templates  $Z_{f,l} \in \mathbb{R}^{N_s \times N_c}$ . The individual templates can be obtained by averaging training data across trials for each visual stimulus. Letting  $S_1$  be  $X$  and  $S_2$  be  $Y_{f,l}$  in (4), we obtain a spatial filter,  $w_1$ , denoted by  $w_{xy}$ . In the same way, we can obtain  $w_{xz}$  from  $X$  and  $Z_{f,l}$  and  $w_{zy}$  from  $Z_{f,l}$  and  $Y_{f,l}$ , respectively. Using these spatial filters, the correlation coefficients between the test data and the templates are computed as:

$$r_{f,l} = \begin{bmatrix} r_{f,l,1} \\ r_{f,l,2} \\ r_{f,l,3} \\ r_{f,l,4} \end{bmatrix} = \begin{bmatrix} \rho_{f,l} \\ \text{corr}(w_{xy}^T X, w_{xy}^T Z_{f,l}) \\ \text{corr}(w_{xz}^T X, w_{xz}^T Z_{f,l}) \\ \text{corr}(w_{zy}^T X, w_{zy}^T Z_{f,l}) \end{bmatrix}, \quad (7)$$

where,  $\text{corr}(\cdot, \cdot)$  is the Pearson's correlation coefficient. The correlation coefficients are combined and used for detecting a target stimulus as follows:

$$f^*, l^* = \arg \max_{f,l} \sum_{i=1}^4 \text{sign}(r_{f,l,i}) \cdot r_{f,l,i}^2 \quad (8)$$

**Table 1.** Accuracy of waveform and/or frequency classification using data length of 2 s for each subject.

Subject	Accuracy [%]		
	Waveform	Frequency	Mixed
s1	96.67	99.44	93.89
s2	88.89	100.00	87.78
s3	97.22	100.00	98.89
s4	86.11	100.00	84.44
s5	92.22	97.22	91.11
s6	98.89	100.00	97.78
Mean (Std)	93.33 (5.10)	99.44 (1.11)	92.31 (5.65)

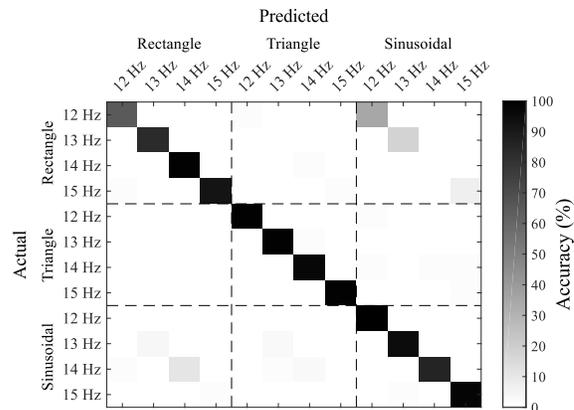
### 3. RESULTS

Fig. 1(b) shows the averaged time series and amplitude spectra of elicited SSVEPs at 12 Hz across trials recorded from s3. Among three stimulus waveforms, the fundamental and harmonic components in SSVEPs showed different amplitude. For example, the amplitude spectrum of the SSVEP elicited by the rectangle waveform had clear peaks at fundamental, second and third harmonics frequencies, which is consistent to that of the stimulus signal. The sinusoidal waveform, on the other hand, elicited less harmonic components than the other modulation waveforms.

Fig. 2 shows the averaged accuracy across six subjects in detecting stimulus waveforms and/or frequencies using the two methods with different data lengths. The classification accuracy was estimated using a leave-one-out cross validation. The CCA-based method achieved higher accuracy than the chance-level in frequency classification when the data length were over 0.4 s. However, regardless of data length, the CCA-based method did not show ability to discriminate the stimulus waveforms. In contrast, the extended CCA-based method achieved significantly higher accuracy than the chance-levels in all three conditions. Table 1 lists the accuracy for each individual when the data length was 2 s. All subjects reached almost perfect accuracy in frequency detection with the averaged accuracy of  $99.44 \pm 1.11\%$ . Although the accuracy of waveform classification was lower than that of frequency classification, four out of six subjects achieved 90% accuracy. Overall averaged accuracy in mixed waveform and frequency classification across subjects was  $92.31 \pm 5.65\%$ .

### 4. DISCUSSIONS

The experimental results showed that the SSVEPs modulated by different waveforms can be classified by using the extended CCA-based method. With data length of 2 s, all the subjects obtained the accuracy over 80% in the classification of waveforms, frequencies, and mixed waveforms and frequencies (Table 1). Interestingly, although two subjects (s2 and s4) reached the perfect accuracy in frequency classification, their accuracy in waveform classification was less than 90%, which was lower than the other subjects. Fig. 3 depicts



**Fig. 3.** Averaged confusion matrix across subjects using the extended CCA-based method.

the averaged confusion matrix derived using the extended CCA-based method. Although most of the trials were classified into its actual classes, there were a few trials misclassified into rectangle waveforms rather than sinusoidal waveforms, and vice versa. These results indicate that the appearance of SSVEP components (i.e., fundamental and harmonic components) is different for each subject even he/she gazes at target stimuli correctly. This may be due to individual differences in transfer functions among photoreceptors in the retina, cortical sources, and the scalp. The waveform classification could be improved by estimating and integrating individual transfer functions into target identification algorithms. These results also suggest the importance of selection and design of waveforms. Systematic design to enhance the classification efficiency would be highly desirable.

As shown in the Table 1, the 12-class SSVEPs can be classified with an averaged accuracy of  $92.31 \pm 5.65\%$  with five electrodes. This accuracy is competitive with the results reported in previous papers [18, 19]. For example, Han et al.'s system, in which 12 visual stimuli were modulated by different spatial patterns, obtained an accuracy of 91.7% with seven electrodes [18]. Xie and Meng's system with 12 visual stimuli modulated by different frequencies selected from a wide frequency range obtained an averaged accuracy of 70.9% with three electrodes [19]. Since the optimal algorithm and electrode setting might be different for each stimulus design, a direct and systematic comparison among stimulus designs would be helpful for future researches. To evaluate the possibility of applying the waveform-based stimulation into an online BCI system, the simulated online ITR [20] was calculated. With 1-s gaze shifting time, the averaged simulated ITR across subjects with data lengths of 1 s and 2 s were  $72.12 \pm 7.35$  bits/min and  $59.16 \pm 8.12$  bits/min, respectively. Importantly, the proposed method can be combined with any types of existing visual stimulation approaches including hybrid frequency and phase coding, leading to a significantly large number of BCI commands with comparable accuracy.

## 5. REFERENCES

- [1] F. B. Vialatte, M. Maurice, J. Dauwels, and A. Cichocki, "Steady-state visually evoked potentials: focus on essential paradigms and future perspectives," *Prog. Neurobiol.*, vol. 90, no. 4, pp. 418–438, Apr. 2010.
- [2] A. M. Norcia, L. G. Appelbaum, J. M. Ales, B. R. Cottereau, and B. Rossion, "The steady-state visual evoked potential in vision research: A review," *J. Vis.*, vol. 15, no. 6, pp. 1–46, 2015.
- [3] S. Gao, Y. Wang, X. Gao, and B. Hong, "Visual and auditory brain-computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 5, pp. 1436–1447, 2014.
- [4] Y. Wang, X. Gao, B. Hong, C. Jia, and S. Gao, "Brain-computer interfaces based on visual evoked potentials," *IEEE Eng. Med. Biol. Mag.*, vol. 27, no. 5, pp. 64–71, 2008.
- [5] D. Regan, "Electrical responses evoked from the human brain," *Sci. Am.*, vol. 24, no. 6, pp. 134–146, 1979.
- [6] M. Nakanishi, Y. Wang, Y. T. Wang, Y. Mitsukura, and T. P. Jung, "A high-speed brain speller using steady-state visual evoked potentials," *Int. J. Neural Eng.*, vol. 24, no. 6, pp. 1450019, 2014.
- [7] X. Chen, Y. Wang, M. Nakanishi, X. Gao, T. P. Jung, and S. Gao, "High-speed spelling with a noninvasive brain-computer interface," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 112, no. 44, pp. E6058–E6067, 2015.
- [8] Y. Wang, R. Wang, X. Gao, B. Hong, and S. Gao, "A practical VEP-based brain-computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 2, pp. 234–239, June 2006.
- [9] M. Nakanishi, Y. Wang, Y. T. Wang, Y. Mitsukura, and T. P. Jung, "Generating visual flickers for eliciting robust steady-state visual evoked potentials at flexible frequencies using monitor refresh rate," *PLoS One*, vol. 9, no. 6, pp. e99235, 2014.
- [10] X. Chen, Z. Chen, S. Gao, and X. Gao, "A high-ITR SSVEP-based BCI speller," *Brain-Comp. Interface*, vol. 1, no. 3-4, pp. 181–191, 2014.
- [11] Y. Zhang, P. Xu, T. Liu, J. Hu, R. Zhang, and D. Yao, "Multiple frequencies sequential coding for SSVEP-based brain-computer interface," *PLoS One*, vol. 7, no. 3, pp. e29519, 2012.
- [12] Y. Kimura, T. Tanaka, H. Higashi, and N. Morikawa, "SSVEP-based brain-computer interfaces using FSK-modulated visual stimuli," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 10, pp. 2831–2838, May 2013.
- [13] C. Jia, X. Gao, B. Hong, and S. Gao, "Frequency and phase mixed coding in SSVEP-based brain-computer interface," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 200–206, 2011.
- [14] F. Teng, Y. Chen, A. M. Choong, S. Gustafson, C. Reichley, P. Lawhead, and D. Waddell, "Square or sine: Finding a waveform with high success rate of eliciting SSVEP," *Comp. Intell. Neurosci.*, vol. 2011, pp. 364385, 2011.
- [15] D. H. Brainard, "The Psychophysics Toolbox," *Spat. Vis.*, vol. 10, no. 4, pp. 433–436, Jan. 1997.
- [16] 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. 54, no. 6, pp. 1172–1176, 2007.
- [17] G. Bin, X. Gao, Y. Wang, B. Hong, and S. Gao, "VEP-based brain-computer interfaces: time, frequency, and code modulations," *IEEE Comput. Intell. Mag.*, vol. 4, no. November, pp. 22–26, 2009.
- [18] C. H. Han and C. H. Im, "Classification of visual stimuli with different spatial patterns for single-frequency multi-class SSVEP BCI," *Electron. Lett.*, vol. 49, no. 22, pp. 1374–1376, 2013.
- [19] S. Xie and W. Meng, *Signal processing methods for SSVEP-based BCIs*, pp. 51–68, Springer International Publishing, Cham, 2017.
- [20] M. Cheng, X. Gao, S. Gao, and D. Xu, "Design and implementation of a brain-computer interface with high transfer rates," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 10, pp. 1181–1186, 2002.