Enhancing Detection of Steady-State Visual Evoked Potentials Using Individual Training Data

Yijun Wang*, *Member, IEEE*, Masaki Nakanishi, *Student Member, IEEE,* Yu-Te Wang, *Student Member, IEEE,* and Tzyy-Ping Jung, *Senior Member, IEEE*

Abstract— **Although the performance of steady-state visual evoked potential (SSVEP)-based brain-computer interfaces (BCIs) has improved gradually in the past decades, it still does not meet the requirement of a high communication speed in many applications. A major challenge is the interference of spontaneous background EEG activities in discriminating SSVEPs. An SSVEP BCI using frequency coding typically does not have a calibration procedure since the frequency of SSVEPs can be recognized by power spectrum density analysis (PSDA). However, the detection rate can be deteriorated by the spontaneous EEG activities within the same frequency range because phase information of SSVEPs is ignored in frequency detection. To address this problem, this study proposed to incorporate individual SSVEP training data into canonical correlation analysis (CCA) to improve the frequency detection of SSVEPs. An eight-class SSVEP dataset recorded from 10 subjects in a simulated online BCI experiment was used for performance evaluation. Compared to the standard CCA method, the proposed method obtained significantly improved detection accuracy (95.2% vs. 88.4%, p<0.05) and information transfer rates (ITR) (104.6 bits/min vs. 89.1 bits/min, p<0.05). The results suggest that the employment of individual SSVEP training data can significantly improve the detection rate and thereby facilitate the implementation of a high-speed BCI.**

I. INTRODUCTION

Steady-state visual evoked potentials (SSVEPs) have been widely used in electroencephalogram (EEG)-based brain-computer interfaces (BCIs) [1, 2]. Frequency coding is the most popular multi-target coding approach in SSVEP BCIs. In such a system, users need to fixate their gaze on one of multiple visual flickers tagged with different stimulation frequencies. The gazed target can be identified through recognizing the frequency of the SSVEPs corresponding to the target stimulus. Currently, an SSVEP BCI typically does not have a calibration procedure since the frequency of SSVEPs can be simply recognized by power spectrum density analysis (PSDA) approaches (e.g., fast Fourier transform, FFT) [3-5]. The use of PSDA-based frequency detection makes it possible to realize a calibration free system. However, since the phase

information of SSVEPs is totally ignored in PSDA-based frequency detection, the detection rate could be seriously deteriorated by the spontaneous EEG activities within the same frequency range of SSVEPs. Although the performance of SSVEP-based BCIs has improved gradually in the past decades, it still does not meet the requirement of a high communication speed in many practical applications [6].

In EEG-based BCIs, spatial filtering techniques have been widely used to improve signal-to-noise ratio (SNR) of EEG signals. Recently, multi-channel decoding methods based on spatial filtering have been applied to SSVEP-based BCIs [7, 8]. Canonical correlation analysis (CCA) is one of the most popular approaches. In an unsupervised way, the standard CCA-based approach finds spatial filters to maximize the correlation between SSVEPs and sine-cosine reference signals [8]. Although CCA can significantly improve the SNR of SSVEPs, the interference from spontaneous EEG activities still exists in the CCA-based frequency detection. Fortunately, the interference problem can be alleviated by characterizing the initial phase in SSVEPs and spontaneous EEG activities. Theoretically, SSVEPs time locked to stimuli at the same frequency have the same initial phase due to a fixed latency delay in the visual system [9], whereas the spontaneous EEG component at the stimulation frequency has a random phase. Therefore, SSVEP training data can be employed to improve frequency detection. Pan et al. [9] proposed a phase-constrained CCA approach by estimating the latency delay in the visual pathway using SSVEP training data. Recently, Zhang et al. proposed multi-way CCA [10] and multi-set CCA [11] approaches that combined SSVEP training data in CCA process. These methods improved the standard CCA approach in complicated ways that are difficult to understand and implement in real practice [9-11]. Compared to these methods, a more feasible solution is to calculate correlation coefficient between single-trial SSVEPs and reference signals obtained from SSVEP training data after CCA-based spatial filtering. The correlation-based template matching approach has been employed in the code modulation VEP-based BCIs and achieved very high BCI performance [12, 13]. However, to our knowledge, the combination of CCA and training data-based template matching has not been reported in SSVEP-based BCIs.

This study proposed to incorporate individual SSVEP training data to improve the CCA-based frequency detection of SSVEPs. In addition to the standard CCA-based method, correlations between the single-trial SSVEPs and SSVEP

reference signals provided complementary information for target identification. To evaluate the performance of the proposed method, this study compared classification performance between the proposed method and the standard CCA-based method using an eight-class SSVEP dataset recorded from 10 subjects. The goal of this study was to develop a more efficient target detection approach for SSVEP-based BCIs, which can facilitate the implementation of a high-speed BCI for practical applications.

II. METHOD

A. Stimulus Presentation

In the conventional frame-based method for rendering visual flickers on a computer monitor, the number of frames in a stimulation cycle is a constant. For example, under a 60Hz refresh rate, a 10Hz stimulus can be realized by reversing the stimulus pattern between black and white every three frames (i.e., six frames per cycle). However, a flickering frequency by which the refresh rate is not dividable (e.g., 11Hz) cannot be presented on the screen using this method. To solve this problem, we proposed an approximation approach that can reliably generate stimulus signals at flexible frequencies [14]. Under a fixed refresh rate, the stimulus sequence $c(f, i)$ for frequency f can be generated by the following equation:

$$c(f, i) = \text{square}\left[2\pi f \left(i/_{RefreshRate}\right)\right] \qquad (1)$$

where square() generates a 50% duty cycle square wave with levels 0 and 1, and i indicates the frame index.

B. Data Acquisition

The dataset was recorded from a simulated online BCI experiment [15]. In the experiment, a Dell S2409W 24-inch LCD monitor (Dell Inc.) with a 75Hz refresh rate was used to present a row of eight flickers (each with a size of 3×3cm) at frequencies from 8Hz to 15Hz with a 1Hz interval. The stimulus sequences were generated by the approximation approach described in (1). The stimulus program was developed under MATLAB (Mathworks Inc.) using the Psychophysics Toolbox extensions [16].

Ten healthy adults (8 males and 2 females, mean age: 23 years) with normal or corrected-to-normal vision participated in the experiment. All subjects signed an informed consent form approved by the Research Ethics Committee of Keio University before participating in the experiment. EEG data were measured using a 16-channel g.USBamp system (g.tec medical engineering GmbH) at a sampling rate of 256 Hz. 12 EEG electrodes (Cz, P1, Pz, P2, PO7, PO3, POz, PO4, PO8, O1, Oz, O2) and four electrooculogram (EOG) electrodes (horizontal and vertical EOG) were placed according to the international 10-20 system.

During the experiment, the subjects were seated in a comfortable chair 70cm away from the monitor in a dark room. They were asked to input a sequence with all eight targets in a task, and to repeat the task 15 times in the experiment. The order of targets was randomized in the task sequence. At the beginning of each trial, a red square marker (3×3cm) appeared at the position of the target stimulus. Subjects were asked to shift their gaze to the target within 0.5 second. At 0.5 second after the cue onset, all stimuli started to flicker concurrently for one second on the monitor. The
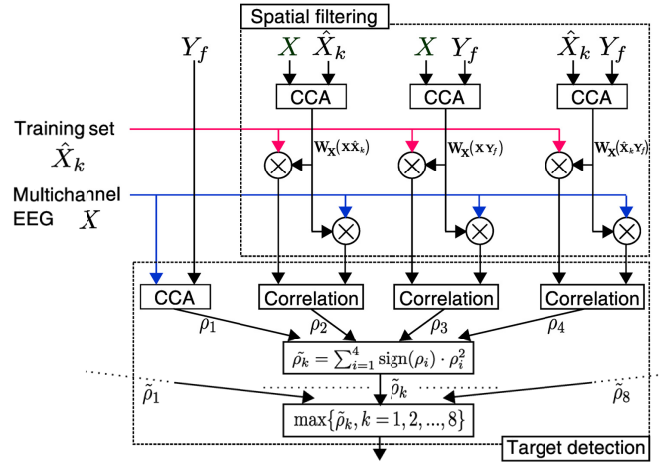

Figure 1. Flowchart of the proposed target identification method.

1s-long EEG data epochs synchronized to the visual stimuli were extracted for target identification.

C. Data Analysis

For each frequency, 15 trials were extracted according to event triggers generated by the stimulus program, resulting in 120 trials for each subject. To avoid overfitting in CCA, eight electrodes (PO7, PO3, POz, PO4, PO8, O1, Oz, O2) over parietal and occipital regions were selected for feature extraction. To remove components irrelevant to SSVEPs, all data epochs were band-pass filtered from 7Hz to 50Hz.

Standard CCA-based approach: In SSVEP-based BCIs, CCA has been widely used to detect the frequency of SSVEPs [8]. CCA is a statistical way used to measure the underlying correlation between two multi-dimensional variables. Considering two multi-dimensional variable X, Y and their linear combinations $x = X^T W_X$ and $y = Y^T W_Y$, CCA finds the weight vectors, W_X and W_Y, which maximize the correlation between x and y by solving the following problem:

$$\max_{W_X, W_Y} \rho(x, y) = \frac{E[W_X^T XY^T W_Y]}{\sqrt{E[W_X^T XX^T W_X]E[W_Y^T YY^T W_Y]}}. \qquad (2)$$

The maximum of ρ with respect to W_X and W_Y is the maximum canonical correlation. In SSVEP detection, X refers to multi-channel SSVEP signals and Y refers to reference signals that have the same length as X. To detect the frequency of SSVEPs in an unsupervised way, sinusoidal signals can be used as the reference signals Y_f:

$$Y_f = \begin{bmatrix} sin(2\pi fn) \\ cos(2\pi fn) \\ \vdots \\ sin(2\pi N_h fn) \\ cos(2\pi N_h fn) \end{bmatrix}, n = \frac{1}{f_s}, \frac{2}{f_s}, \cdots, \frac{N}{f_s} \qquad (3)$$

where f is the target frequency, N_h is the number of harmonics ($N_h = 3$), and N is the number of sampling points. To recognize the frequency of SSVEPs, CCA calculates the canonical correlation between multi-channel SSVEP signals and the reference signals at each stimulation frequency. The frequency of the reference signals with the maximal correlation is considered as the frequency of SSVEPs.

CCA with SSVEP training data: This study developed an extended CCA-based method to incorporate SSVEP training data in target identification. Fig. 1 shows the flowchart of the proposed method. In addition to the standard CCA method, the proposed method combined correlations between single-trial SSVEPs and SSVEP reference signals in target identification. For each stimulation frequency, the training SSVEP reference signals $\widehat{X}$ can be obtained by averaging multiple SSVEP trials in a training set. Correlation coefficients between projections of test set X and training reference signals $\widehat{X}$ using CCA-based spatial filters can be used as features. Specifically, the following three weight vectors are used as spatial filters to enhance the SNR of SSVEPs: (1) $W_X(X\widehat{X})$ between test set X and training reference signals $\widehat{X}$, (2) $W_X(XY)$ between test set X and sine-cosine reference signals Y, and (3) $W_X(\widehat{X}Y)$ between training reference signals $\widehat{X}$ and sine-cosine reference signals Y. The target can be identified by recognizing the training reference signal that maximizes the correlation to X. A correlation vector ρ is defined as follows:

$$\boldsymbol{\rho} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{bmatrix} = \begin{bmatrix} \rho(X^T W_X(XY), Y^T W_Y(XY)) \\ \rho(X^T W_X(X\widehat{X}), \widehat{X}^T W_X(X\widehat{X})) \\ \rho(X^T W_X(XY), \widehat{X}^T W_X(XY)) \\ \rho(X^T W_X(\widehat{X}Y), \widehat{X}^T W_X(\widehat{X}Y)) \end{bmatrix} \quad (4)$$

where $\rho(a, b)$ indicates the correlation coefficient between a and b. An ensemble classifier can be used to combine features derived from the four methods described above. In practice, the following weighted correlation coefficient $\tilde{\rho}$ is used as the final feature in target identification:

$$\tilde{\rho} = \sum_{i=1}^{4} \text{sign}(\rho_i) \cdot \rho_i^2 \quad (5)$$

where $\text{sign}()$ is used to remain discriminative information from negative correlation coefficients between test set X and training reference signals $\widehat{X}$. The training reference signal that maximizes the weighted correlation value is selected as the reference signal corresponding to the target.

D. Performance Evaluation

Individual SSVEP training data were required to implement the proposed method. This study used a leave-one-out cross-validation to estimate simulated BCI performance. Training SSVEP reference signals were obtained from the training data in cross-validation. Classification accuracy and information transfer rate (ITR) [15] were calculated for the proposed method and the standard CCA method separately. To estimate the optimal BCI performance, this study also calculated accuracy and ITR using different data lengths.

III. RESULTS

Fig. 2 shows an example of correlation values for a 10Hz SSVEP trial in target identification. Due to the interference from spontaneous EEG, the standard CCA method obtained higher correlation values at 8Hz and 12Hz than 10Hz, resulting in an error in target detection. In contrast, the proposed method obtained a distinct peak value at 10Hz. The correlation values at 8Hz and 12Hz were significantly decreased after combining correlation values derived from correlation analysis with SSVEP training data. Fig. 3 illustrates averaged correlation values for all stimulation
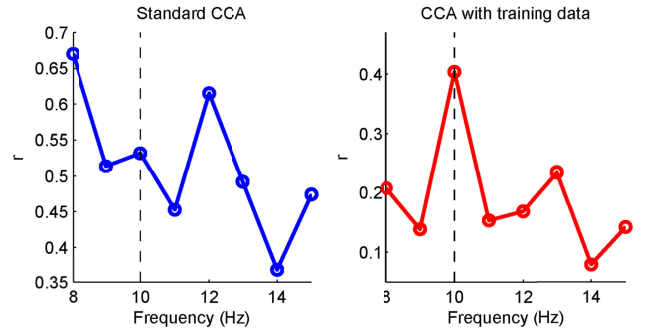

Figure 2. Correlation values between a 10Hz SSVEP trial and reference signals at all stimulation frequencies (8-15Hz).
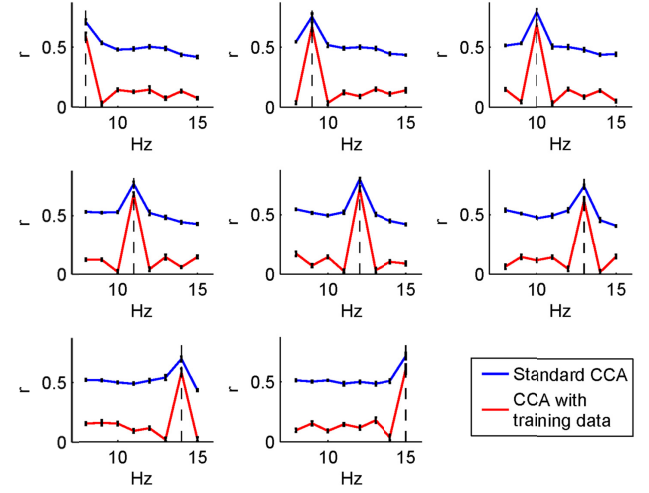

Figure 3. Averaged correlation values between SSVEP trials and reference signals at all stimulation frequencies. The dotted line in each subfigure indicates the stimulation frequency.

frequencies across all subjects. The proposed method significantly decreased the correlation values at non-target frequencies, while the value at the stimulation frequency remained at the same level with a minor decrease. Therefore, the proposed method significantly enhanced the discriminability between target SSVEPs and non-target background EEG activities.

Table I lists the classification accuracy and simulated online ITR using the standard CCA-based method and the proposed method. Paired t-tests showed that the proposed approach obtained significantly higher classification accuracy (95.2% vs. 88.4%, p<0.05) and ITR (104.6 bits/min vs. 89.1 bits/min, p<0.05) than the standard CCA-based method. The subjects with lower performance (<90% using the standard CCA method) achieved more significant improvements using the proposed method (S1: 10.0%, S3: 15.0%, S4: 10.0%, S10: 22.5%). To our knowledge, the resulting ITR of 104.6 bits/min was higher than ITRs reported in other SSVEP-based BCIs using the frequency coding method [1, 2].

Fig. 4 illustrates classification accuracy and ITR using different data lengths from 0.1s to 1s with a 0.1s step. For all conditions, the proposed method obtained significantly improved performance. The difference was more significant when using a short data length. The proposed method obtained a highest ITR of 132.0 bits/min with a 0.4s data length. Paired t-tests indicated significant improvements of accuracy and

TABLE I. SIMULATED ONLINE BCI PERFORMANCE

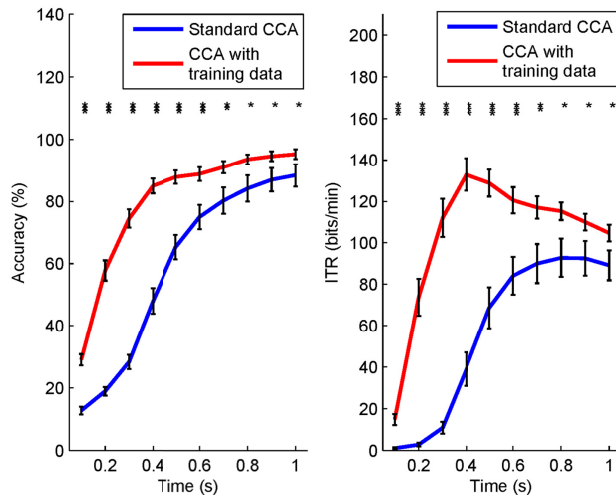| Subject	Accuracy (%)		ITR (bits/min)	
	Standard CCA	CCA with training data	Standard CCA	CCA with training data
S1	88.3	98.3	86.1	113.2
S2	93.3	98.3	98.3	113.2
S3	83.3	98.3	75.2	113.2
S4	85.8	95.8	80.5	105.3
S5	91.6	95.8	94.0	105.3
S6	99.1	97.5	116.2	110.4
S7	93.3	95.8	98.3	105.3
S8	91.6	93.3	94.0	98.3
S9	98.3	97.5	113.2	110.4
S10	59.1	81.6	35.1	71.9
Mean	88.4±11.4	95.2±5.0	89.1±22.9	104.6±12.4

Figure 4. Averaged classification accuracy and ITR using different data lengths. Error bars indicate standard errors. The asterisks indicate significant difference between the two methods (*: p<0.05, **: p<0.01, and ***: p<0.001).

ITR (Accuracy: 47.9±13.1% vs. 85.0±7.6%, p<10^{-6}; ITR: 39.0±25.6 bits/min vs. 132.0±25.0 bits/min, p<10^{-6}).

IV. CONCLUSIONS AND DISCUSSIONS

To solve the problem of interference from spontaneous background EEG activities in SSVEP discrimination, this study proposed to incorporate individual SSVEP training data to improve the detection of SSVEPs. In addition to the standard CCA-based method, correlations between single-trial SSVEPs and SSVEP reference signals were used as new features for target identification. In correlation analysis, this study designed three CCA-based spatial filters to improve the SNR of SSVEPs. The proposed method significantly improved BCI performance over the standard CCA-based method in terms of accuracy (95.2% vs. 88.4%, p<0.05) and ITR (104.6 bits/min vs. 89.1 bits/min, p<0.05) on a dataset recorded from an SSVEP BCI with a high selection speed (1.5 seconds per selection). The results using different data lengths further suggested that the proposed method could be especially useful for high-speed BCIs that can select a target within several hundred milliseconds (see Fig. 4).

The SSVEP dataset used in this study was from a gaze dependent BCI [6]. The proposed method is also applicable to gaze independent BCIs based on visual selective attention [17]. In addition to frequency detection of SSVEPs, the proposed method can also benefit detection of SSVEPs coded by other multi-target coding methods [6]. For example, individual training data were typically required for target identification in BCIs using phase coding [1] or mixed frequency and phase coding [18]. Therefore, the proposed method can be easily adopted to improve target detection in these systems without extra effort in data collection. Furthermore, other future work will include optimization of the ensemble classifier illustrated in (5) and a quantitative comparison of performance between the proposed method and other extended CCA-based methods [9-11].

REFERENCES

[1] Y. Wang, X. Gao, B. Hong, C. Jia, and S. Gao, "Brain-computer interfaces based on visual evoked potentials: feasibility of practical system design," *IEEE EMB. Mag.*, vol. 27, no. 5, pp. 64–71, 2008.

[2] G. Bin, X. Gao, Y. Wang, B. Hong, and S. Gao, "Research frontier: VEP-based brain-computer interfaces: time, frequency, and code modulations," *IEEE Comput. Intell. Mag.*, vol. 4, no. 4, pp. 22–26, 2009.

[3] M. Middendorf, G. McMillan, G. Calhoun, and K. S. Jones, "Brain-computer interfaces based on the steady-state visual-evoked response," *IEEE Trans. Rehab. Eng.*, vol. 8, no. 2, pp. 211–214, 2000.

[4] 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.

[5] Y. Wang, R. Wang, X. Gao, B. Hong, and S. Gao, "A practical VEP-based brain-computer interface," *IEEE Trans. Neural Syst. Rehab. Eng.*, vol. 14, no. 2, pp. 234–240, 2006.

[6] 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.

[7] O. Friman, I. Volosyak, and A. Graser, "Multiple channel detection of steady-state visual evoked potentials for brain-computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 4, pp. 742–750, 2007.

[8] 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.

[9] J. Pan, X. Gao, F. Duan, Z. Yan and S. Gao, "Enhancing the classification accuracy of steady-state visual evoked potential-based brain-computer interfaces using phase constrained canonical correlation analysis," *J. Neural Eng.*, vol. 8, no. 3, p. 036027, 2011.

[10] Y. Zhang, G. Zhou, J. Jin, M. Wang, X. Wang, and A. Cichocki, "L1-Regularized Multiway Canonical Correlation Analysis for SSVEP-Based BCI," *IEEE Trans. Neural Syst. Rehab. Eng.*, vol. 21, no.6, pp. 887–896, 2013.

[11] 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. 2, pp. 1–14, 2014.

[12] E. E. Sutter, "The brain response interface: communication through visually-induced electrical brain responses," *J. Microcomput. Appl.*, vol. 15, pp. 31–45, 1992.

[13] G. Bin, X. Gao, Y. Wang, Y. Li, B. Hong, and S. Gao, "A high-speed BCI based on code modulation VEP," *J. Neural Eng.*, vol. 8, no. 2, p. 025015, 2011.

[14] Y. Wang, Y.-T. Wang, and T.-P. Jung, "Visual stimulus design for high-rate SSVEP," *Electron. Lett.*, vol. 46, no. 15, pp. 1057–1058, 2010.

[15] 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, e99235, 2014.

[16] D. H. Brainard, "The psychophysics toolbox," *Spat. Vis.*, vol. 10, no. 4, pp. 433–436, 1997.

[17] S. P. Kelly, E. C. Lalor, C. Finucane, G. McDarby and R. B. Reilly, "Visual spatial attention control in an independent brain-computer interface," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 9, pp. 1588–1596, 2005.

[18] 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. 12, pp. 200–206, 2011.