Enhancing Unsupervised Canonical Correlation Analysis-Based Frequency Detection of SSVEPs by Incorporating Background EEG

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

Abstract— Steady-state visual evoked potential (SSVEP)-based brain-computer interfaces (BCIs) have potential to provide a fast communication channel between human brain and external devices. In SSVEP-based BCIs, Canonical Correlation Analysis (CCA) has been widely used to detect frequency-coded SSVEPs due to its high efficiency and robustness. However, the detectability of SSVEPs differs among frequencies due to a power-law distribution of the power spectra of spontaneous electroencephalogram (EEG) signals. This study proposed a new method based on the fact that changes of canonical correlation coefficients for SSVEPs and background EEG signals follow the same trend along frequency. The proposed method defined a normalized canonical correlation coefficient, the ratio of the canonical correlation coefficient for SSVEPs to the mean of the canonical correlation coefficients for background EEG signals, to enhance the frequency detection of SSVEPs. An SSVEP dataset from 13 subjects was used for comparing classification performance between the proposed method and the standard CCA method. Classification accuracy and simulated information transfer rates (ITR) suggest that, in an unsupervised way, the proposed method could considerably improve the frequency detection accuracy of SSVEPs with little computational effort.

I. INTRODUCTION

Steady-state visual evoked potentials (SSVEPs) are brain's electrical responses to repetitive visual stimulation, which have sinusoidal-like waveforms at the stimulation frequency and its harmonics. SSVEP has been widely used in electroencephalogram (EEG)-based brain-computer interfaces (BCIs) since it has many advantages such as little user training, ease of use, and high information transfer rate (ITR) [1]–[3]. In an SSVEP-based BCI, users are asked to gaze at one of multiple visual stimuli flickering at different frequencies. The target stimulus, which a user is gazing at, can be identified through frequency detection by analyzing the SSVEPs in frequency domain. As a result, an SSVEP-based BCI can directly translate intentional brain activities into commands tagged with stimulation frequencies to control an external device.

An accurate and fast method for frequency detection plays an important role in SSVEP-based BCIs [3]. Canonical correlation coefficients between SSVEPs and sinusoidal templates (i.e. sines and cosines at the stimulation frequency and its higher harmonic frequencies) calculated by canonical correlation analysis (CCA) have been successfully used in SSVEP-based BCIs to detect the frequency of SSVEPs [4], [5]. Compared to the power spectrum density (PSD)-based frequency detection approach (e.g., fast Fourier transform (FFT)), the CCA-based approach significantly improved the detection accuracy. However, the standard CCA-based frequency detection approach considers each frequency in the same way, leading to different detection accuracies across stimulation frequencies. For example, Lin et al. reported that the coefficients tended to decrease as stimulation frequency increased, resulting in decreased accuracy of SSVEP detection in high stimulation frequencies [6]. Several recent studies have tried to alleviate this problem. Tanaka et al. proposed a machine learning-based method, which combined CCA and linear discriminant analysis (LDA), to improve the detection of SSVEPs in the higher frequency range [7]. Nevertheless, the training procedure increases the difficulty of system use in real-life applications. In another study, Wei et al. proposed a differential CCA (dCCA) method, which considered the frequency that showed a highest degrade of canonical correlation between SSVEPs and notch-filtered SSVEP signals as stimulation frequency [8]. Although this method could enhance the detectability of high-frequency SSVEPs, it is not applicable to epochs with short duration since SSVEP signals can be deteriorated by the transient response of a notch filter.

Wang et al. reported that the signal-to-noise ratio (SNR) of SSVEPs among low-, medium- and high-frequency regions could have comparable levels despite the fact that amplitude of SSVEPs decreased while the stimulation frequency increased across the three frequency regions [9]. In the CCA-based analysis, a normalization procedure similar to the calculation of SNR could be useful to improve the detection of SSVEPs in the higher frequency bands. This study therefore proposes a new unsupervised method to derive normalized canonical correlation coefficients, which are defined as canonical correlation coefficients for SSVEPs divided by the mean of canonical correlation coefficients for neighboring background EEG signals, to improve the frequency detection of SSVEPs.
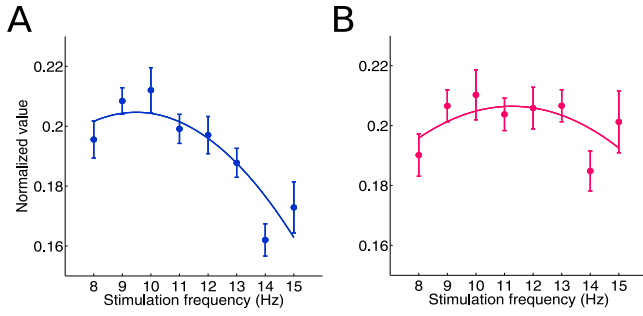
Fig. 1. (A) The canonical correlation coefficients $r(f)$ and (B) the normalized coefficients $\tilde{r}(f)$ for SSVEP epochs at different stimulation frequencies. Solid lines indicate regression curves. Error bars indicate standard errors.

II. METHODS

A. EEG Acquisition

The EEG data were recorded in an SSVEP BCI study using the mixed frequency and phase coding method [10]. In the experiment, 32 visual flickering stimuli (3 × 3 cm squares) were concurrently rendered on a 24-inch liquid crystal display (LCD) with a refresh rate of 75Hz. The 32 stimuli in a 4 × 8 matrix were specified by four phases (0, 90, 180, 270 degrees) and eight frequencies (8Hz to 15Hz with an interval of 1Hz). The horizontal and vertical intervals between two stimuli were 2 cm and 3 cm respectively. The stimulation frequencies and phases were generated using an approximation approach described in [10]. The stimulation program was developed under MATLAB (Mathworks, Inc.) using the Psychophysics Toolbox extensions [11]. This study only focused on frequency detection (eight frequencies from 8 to 15Hz).

Thirteen healthy adults (10 males and 3 females, mean age: 23 years) with normal or corrected-to-normalized 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. The subjects were seated in a comfortable chair 70 cm in front of the LCD in a dark room. They were asked to gaze at one of the visual stimuli (a target stimulus) for 4 seconds, and 32 targets were indicated in a random order in a run. There was a time interval of 0.5 second between two consecutive trials. Subjects were asked to shift their gaze to the next target within the 0.5 s duration. Seven runs were carried out for each subject. EEG data were recorded by 16 electrodes covering the parietal and occipital areas (FPz, F3, F4, Fz, Cz, P1, P2, Pz, PO3, PO4, PO7, PO8, POz, O1, O2 and Oz) using the g.USBamp (g.tec medical engineering GmbH) with a sampling rate of 512Hz.

B. EEG Analysis

1) Standard CCA-based method: This study first performed offline classification of SSVEPs at all eight frequencies using a standard CCA-based method [4], [5]. CCA is a statistical method used to measure underlying correlation between two multidimensional variables. Considering two multidimensional variables $\mathbf{X}$, $\mathbf{Y}$ and their linear combinations $\mathbf{x} = \mathbf{X}^T\mathbf{W_x}$ and $\mathbf{y} = \mathbf{Y}^T\mathbf{W_y}$, CCA finds the weight vectors, $\mathbf{W_x}$ and $\mathbf{W_y}$, which maximize the correlation between $\mathbf{x}$ and $\mathbf{y}$ by solving the following problem:

$$\max_{\mathbf{W_x},\mathbf{W_y}} r = \frac{E[\mathbf{W_x}^T\mathbf{X}\mathbf{Y}^T\mathbf{W_y}]}{\sqrt{E[\mathbf{W_x}^T\mathbf{X}\mathbf{X}^T\mathbf{W_x}]E[\mathbf{W_y}^T\mathbf{Y}\mathbf{Y}^T\mathbf{W_y}]}}. \quad (1)$$

The maximum of r with respect to $\mathbf{W_x}$ and $\mathbf{W_y}$ is the maximum canonical correlation. Here, $\mathbf{X}$ refers to the set of multi-channel EEG signals and $\mathbf{Y}$ refers to the set of reference signals that have the same length as $\mathbf{X}$. The reference signals $\mathbf{Y}_f$ are set as

$$\mathbf{Y}_f = \begin{bmatrix} sin(2\pi ft) \\ cos(2\pi ft) \\ \vdots \\ sin(2\pi N_h ft) \\ cos(2\pi N_h ft) \end{bmatrix} \quad (2)$$

where, f is the target frequency and N_h is the number of harmonics. To detect the frequency of SSVEPs, CCA calculates the canonical correlation between multi-channel EEG signals and the reference signals at each stimulus
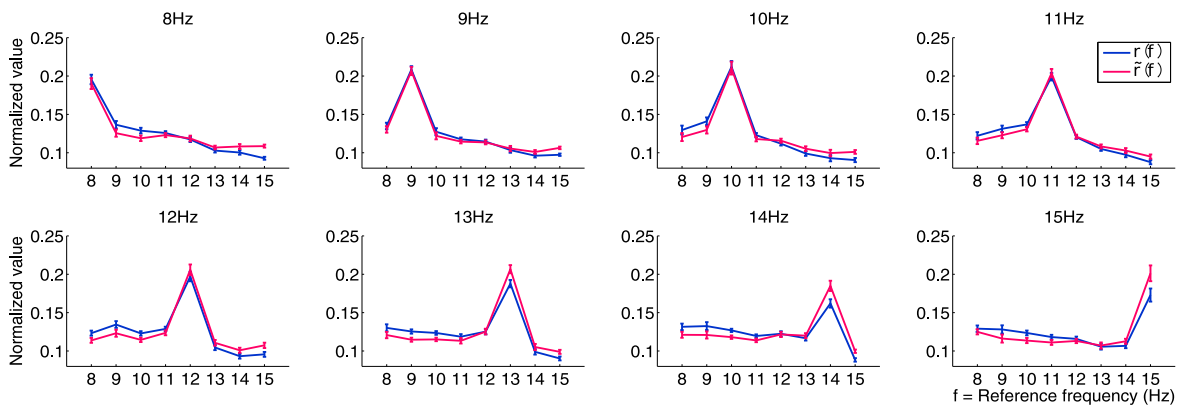

Fig. 2. A comparison between averaged $r(f)$ and $\tilde{r}(f)$ across all subjects for different stimulation frequencies (8-15Hz) with 1s-long data. The $\tilde{r}(f)$ were calculated with N_h=2, Δf=1.0Hz and K=6. The values were normalized so that $\sum r(f) = 1$ and $\sum \tilde{r}(f) = 1$. Error bars indicate standard errors.
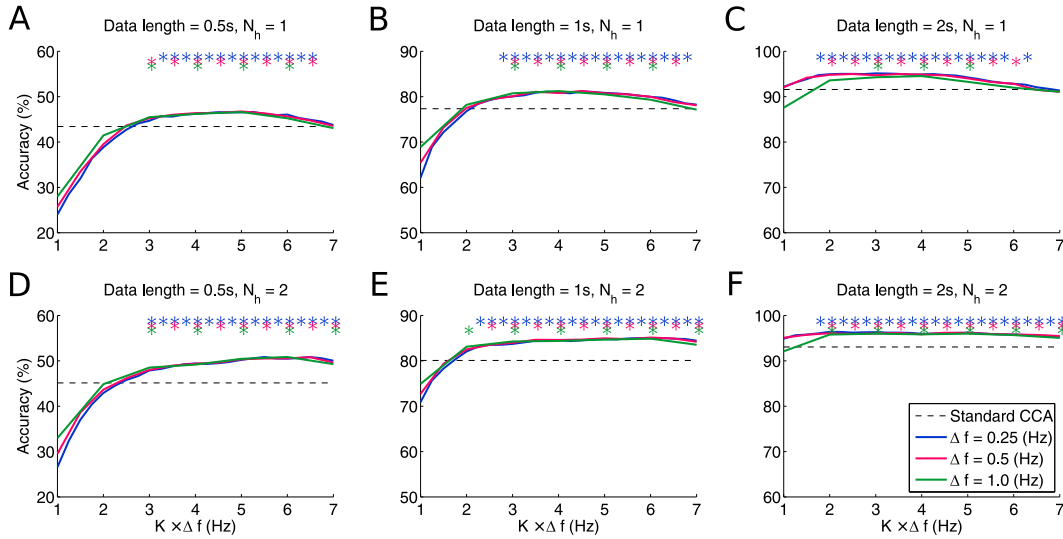
Fig. 3. Averaged classification accuracy across subjects calculated by the standard CCA-based method and the proposed method with different data length, the number of harmonics, Δf and K. (A)-(C) The accuracy with fundamental frequency (N_h=1). (D)-(F) The accuracy with fundamental and second harmonics (N_h=2). The asterisks indicate significant differences between two methods tested by one-tail t-test (*p<0.05).

frequency. The frequency of the reference signals with the maximal canonical correlation was selected as the SSVEP frequency.

2) The proposed method: The detectability of the standard CCA-based method in higher frequency range is low since the canonical correlation between multichannel EEG signal and the reference signals tends to decrease as the stimulation frequency increases. This study defines the ratio of canonical correlation coefficient for SSVEPs to those for neighboring background EEG signals as normalized canonical correlation for frequecy detection. For each stimulation frequency f, normalized canonical correlation value $\tilde{r}(f)$ can be calculated by deviding the canonical correlation for the SSVEPs $r(f)$ by the mean of canonical correlations for the background EEG activities within the neighboring frequency bands $r(f \pm k\Delta f)$:

$$\tilde{r}(f) = \frac{K \times r(f)}{\sum_{k=1}^{K}\left(r(f+k\Delta f)+r(f-k\Delta f)\right)} \quad (3)$$

where, f is stimulation frequency, Δf is frequency resolution, and K is the number of neighboring frequencies on each side. The frequency of the reference signals f with maximal $\tilde{r}(f)$ was selected as the stimulation frequency.

C. Performance Evaluation

The recorded EEG data were first down-sampled to 256Hz, and divided into 4s-long EEG epochs according to event triggers generated by the stimulus program. All data epochs were referenced to a common average reference (CAR) and then band-pass filtered from 5Hz to 40Hz with an infinite impulse response (IIR) filter. To avoid overfitting in CCA, four electrodes over the occipital region (POz, O1, O2, and Oz) were selected for representing the SSVEP data. Two methods including the standard CCA-based method and the proposed method were used to estimate the frequency detection accuracy. In addition to classification accuracy,

BCI performance was also evaluated by ITR calculated as follows:

$$ITR = \frac{60}{T}\left\{\log_2 M + P\log_2 P + (1-P)\log_2\left(\frac{1-P}{M-1}\right)\right\} \quad (4)$$

where, M is the number of targets, P is the classification accuracy, and T (seconds/selection) is the average time for a selection (including target gazing time and gaze shifting time). This study calculated classification performance using different T (Target gazing: 0.5s, 1s, 2s; Gaze shifting: 0.5s).

III. RESULTS

Fig. 1 shows the normalized value of $r(f)$ and $\tilde{r}(f)$ for SSVEPs at the stimulation frequency with 1s-long epochs (N_h=2, Δf =1.0 Hz, and K=6). Fig. 2 shows the normalized value of $r(f)$ and $\tilde{r}(f)$ for each stimulation frequency. Fig. 1A shows that the values of canonical correlation $r(f)$ differed across stimulation frequencies. The values at 14Hz and 15Hz, for example, were significantly lower than those at other frequencies. In contrast, the $\tilde{r}(f)$ for all stimulation frequencies show comparable values (Fig. 1B). As shown in Fig. 2, compared to $r(f)$, the value of $\tilde{r}(f)$ at the stimulation frequency was higher at 12-15Hz, while it remained at the same level at other frequencies. These results suggest that the proposed method can effectively improve the detection of frequencies in the higher frequency range.

Fig. 3 shows the averaged accuracy for the eight-frequency classification across all subjects with different parameter settings (epoch length, N_h, Δf, and K). The results indicate that the classification accuracy with three different Δf (0.25Hz, 0.5Hz, and 1.0Hz) were comparable to each other. In all conditions, the proposed method obtained significantly higher accuracy than the standard CCA-based method. For each condition, the classification accuracy increased as the bandwidth of neighboring frequencies ($K \times \Delta f$) increased till it reached the highest value with a bandwidth of 6Hz (N_h=2).

TABLE I

OFFLINE SSVEP DETECTION ACCURACY AND ITR

Subject	Accuracy (%)		ITR (bits/min)	
	Standard	Proposed	Standard	Proposed
s1	91.96	92.41	94.84	95.98
s2	75.89	80.36	61.06	69.35
s3	85.27	91.07	79.33	92.61
s4	59.82	66.07	36.00	44.93
s5	85.71	92.41	80.29	95.98
s6	78.12	88.84	65.12	87.28
s7	87.50	91.52	84.22	93.72
s8	95.09	96.88	103.18	108.47
s9	41.96	54.91	15.58	29.65
s10	89.73	93.30	89.37	98.30
s11	79.46	84.38	67.64	77.44
s12	92.86	93.75	97.13	99.49
s13	77.68	77.68	64.29	64.29
Mean	80.08	84.89	72.16	81.35
STD	14.80	12.35	24.87	23.35

*Data length = 1s, N_h = 2, Δf = 1.0Hz, K = 6

TABLE II

DETECTION ACCURACY (%) FOR EACH STIMULATION FREQUENCY

Stimulation frequency	Standard	Proposed	Difference
8Hz	83.79	81.87	-1.92
9Hz	90.66	87.91	-2.75
10Hz	90.11	87.64	-2.47
11Hz	87.36	87.64	0.28
12Hz	83.52	89.29	5.77
13Hz	77.47	85.99	8.52
14Hz	58.24	78.85	20.61
15Hz	69.51	79.95	10.44

*Data length = 1s, N_h = 2, Δf = 1.0Hz, K = 6

Table I lists the classification accuracy for each subject using the standard CCA-based method and the proposed method with 1s epochs (Δf=1.0Hz, K=6, N_h=2). The averaged accuracy obtained by the standard CCA-based method and the proposed method were 80.08±14.80% and 84.89±12.35%, respectively. The difference between the two methods was significant (paired t-test, p<0.01). The estimated ITR of the proposed method was significantly higher than that of the standard CCA-based method (81.35±23.35 vs. 72.16±24.87 bits/min, p<0.001).

IV. DISCUSSIONS

Although CCA has been widely used for frequency detetion in SSVEP-based BCIs, lower detectability of SSVEPs at higher stimulation frequency according to a power-law distribution of EEG spectra has remained an issue. By incorporating the background EEG into the standard CCA-based method, this study successfully improved the detection of SSVEPs in the higher frequency range. As shown in Fig. 2, the values of $\tilde{r}(f)$ at the stimulation frequency in the proposed method were higher than those of $r(f)$ in the standard CCA-based method in the higher frequency range (12Hz-15Hz). This result implies that the proposed method can improve the detection accuracy of higher stimulation frequencies. Table II lists classification accuracy for each stimulation frequency. The proposed method improved the classification accuracy for the higher frequencies (11-15Hz).

However, classification performance at the lower frequencies (8-10Hz) decreased at the same time. The overall classification performance significantly increased using the proposed method.

This study compared the classification performance between the standard CCA-based method and the proposed method with different data length, N_h, Δf, and K. In all conditions using different data lengths and N_h, the proposed method obtained significant improvements of the averaged accuracy across subjects in most cases where $3 \leq K \times \Delta f \leq 6$ (Fig. 3), which suggests Δf and K should be selected within these values for any subject. With 1s data length, N_h=2 and $K \times \Delta f$=6, the proposed method substantially outperformed the standard CCA-based method in classification accuracy (84.89% vs. 80.08%, p<0.01) and ITR (81.35 bits/min vs. 72.16 bits/min, p<0.001). Importantly, since the improvements of the accuracy were more notable with shoft data length, the propsoed approach could be especially useful for a BCI system that requires a high selection speed (e.g., 1.5s per selection).

Although this study focused only on the frequency range around EEG alpha band (8-13Hz), the proposed method would be applicable for other frequency bands above critical flicker fusion (CFF) frequencies to reduce visual fatigue [6]. The main concern in our future work will be to apply the proposed method to the implementation of a multi-command, real-time, portable, and visual fatigue-free BCI system.

REFERENCES

[1] 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.

[2] 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.

[3] 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.

[4] 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.

[5] G. Bin, X. Gao, Z. Yan, B. Hong, and S. Gao, "An online multi-channel SSVEP-based brain-computer interface using a canonical correlation analysis method," *J. Neural Eng.*, vol. 6, no. 4, p. 046002, Aug. 2009.

[6] F.-C. Lin, J. K. Zao, K.-C. Tu, Y. Wang, Y.-P. Huang, C.-W. Chuang, H.-Y. Kuo, Y.-Y. Chien, C.-C. Chou, and T.-P. Jung, "SNR analysis of high-frequency steady-state visual evoked potentials from the foveal and extrafoveal regions of human retina," in *34th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Jan. 2012, pp. 1810–1814.

[7] T. Tanaka, C. Zhang, and H. Higashi, "SSVEP frequency detection methods considering background EEG," in *2012 Jt. 6th Int. Conf. Soft Comput. Intell. Syst. 13th Int. Symp. Adv. Intel. Syst.*, 2012, pp. 1138–1143.

[8] C.-S. Wei, Y.-P. Lin, Y. Wang, Y.-T. Wang, and T.-P. Jung, "Detection of steady-state visual-evoked potential using differential canonical correlation analysis," *6th Annu. Int. IEEE EMBS Conf. Neural Eng.*, pp. 57–60, 2013.

[9] 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, Jun. 2006.

[10] 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. Syst.*, 2014, (Accepted).

[11] D. H. Brainard, "The Psychophysics Toolbox," *Spat. Vis.*, vol. 10, no. 4, pp. 433–436, Jan. 1997.