Detection of Steady-state Visual-evoked Potential Using Differential Canonical Correlation Analysis

Chun-Shu Wei, Yuan-Pin Lin, *Member, IEEE*, Yijun Wang, *Member, IEEE*, Yu-Te Wang, *Student Member, IEEE*, and Tzyy-Ping Jung, *Senior Member, IEEE*

***Abstract*—Steady-state visual evoked potential (SSVEP) is an electroencephalogram (EEG) activity elicited by periodic visual flickers. Frequency-coded SSVEP has been commonly adopted for functioning brain-computer interfaces (BCIs). Up to date, canonical correlation analysis (CCA), a multivariate statistical method, is considered to be state-of-the-art to robustly detect SSVEPs. However, the spectra of EEG signals often have a 1/f power-law distribution across frequencies, which inherently confines the CCA efficiency in discriminating between high-frequency SSVEPs and low-frequency background EEG activities. This study proposes a new SSVEP detection method, differential canonical correlation analysis (dCCA), by incorporating CCA with a notch-filtering procedure, to alleviate the frequency-dependent bias. The proposed dCCA approach significantly outperformed the standard CCA approach by around 6% in classifying SSVEPs at five frequencies (9-13Hz). This study could promote the development of high-performance SSVEP-based BCI systems.**

I. Introduction

Steady-state visual evoked potential (SSVEP) is a periodic electroencephalogram (EEG) activity elicited by a flickering stimulus. For the last decade, SSVEP has been intensively adopted in brain-computer interfaces (BCIs) to bridge the human brain with computers or external devices [1-4]. For example, Wang et al. [1] recently demonstrated the feasibility of using a mobile SSVEP-based BCI platform to make a phone call. The performance of such a practical BCI application substantially depends on the detectability of SSVEPs. Accordingly, how to rapidly and accurately decode the frequency-tagged SSVEPs plays an important role in the SSVEP-based BCIs.

Canonical correlation analysis (CCA) has been widely adopted in SSVEP-based BCIs to obtain robust SSVEP detectability [3, 4]. Lin et al. [5] specifically reported that the canonical correlation value, i.e. the correlation between SSVEP and the stimulating signal, tended to decrease as flickering frequency increased. Wang et al. [6] reported that the signal-to-noise ratio (SNR) of SSVEP is frequency-dependent and prone to degrade along ascending frequencies [6]. These might be attributed to the fact that the spectra of EEG signals often have a 1/f power-law distribution across frequencies, leading to degraded magnitude in higher frequencies. The attenuated EEG activity can significantly hinder the detection of SSVEPs in relatively high frequencies. To solve this frequency bias, Tanaka et al. [7] recently proposed to incorporate CCA with linear discriminative analysis (LDA), which is a widely used pattern-recognition method, to improve the detection of SSVEPs in high frequency range. Nevertheless, the training process involved in the recognition strategy could hinder the practicality of SSVEP-based BCIs in real-life applications.

This study addresses the problem of SSVEP detectability across different frequencies caused by the power-law spectral bias. An extension of the conventional CCA method, differential CCA (dCCA), is proposed for improving the detectability of SSVEPs in relatively high frequencies. Compared to the machine learning-based method, the proposed training-free method could be more generalized and practical for online BCI applications.

II. Materials and Methods

A. Experiment Settings

The visual stimulus was delivered on a 21" CRT monitor with a 120Hz refresh rate. A 5×5 cm^2 square in the center of the screen flickered at the frequencies ranged from 9Hz to 13Hz with an interval of 1Hz. The flickering frequency was approximated according to the approach proposed by Wang in 2010 [8]. The system was programmed by Microsoft Visual C++ using the Microsoft DirectX 9.0 framework and operated on Windows XP platform.

During the experiment, subjects sat still in front of the monitor with a fixed distance of 35cm. This experiment used a fixed chin rest to hold a subject's head to avoid head movements. The subjects were instructed to gaze at the frequency-coded visual flickers presented at the monitor in random order. The experiment included four sessions. Each session consisted of five 30s-long blocks corresponding to five frequencies (9-13Hz). A ~15-s rest was interleaved with stimulating blocks for preventing visual fatigue. A minute(s)-long break was also provided between two consecutive sessions.

B. Data Acquisition and Pre-processing

Twelve healthy male subjects with normal or corrected-to-normal vision participated in this study. All participants were asked to read and sign an informed consent form approved by the UCSD Human Research Protections

*Research is supported in part by US Office of Naval Research, Army Research Office (W911NF-09-1-0510), Army Research Laboratory (W911NF-10-2-0022), and DARPA (DARPA/USDI D11PC20183).

Chun-Shu Wei, Yuan-Pin Lin, Yijun Wang, Yu-Te Wang, and Tzyy-Ping Jung are with the Swartz Center of Computational Neuroscience, University of California, San Diego (UCSD), La Jolla, CA, USA.

Chun-Shu Wei and Tzyy-Ping Jung are also with the Department of Bioengineering, University of California, San Diego (UCSD), La Jolla, CA, USA. (email: jung@sccn.ucsd.edu; phone:858-822-7555; fax: 858-822-7556).

Program before the experiment. The data from two subjects were excluded from further analysis due to poor signal quality.

EEG data were recorded using Ag/AgCl electrodes of a BioSemi ActiveTwo EEG system (Biosemi, Inc.) with 256 channels. A 3-D digitizer system (Polhemus, Inc.) was adopted to digitize the electrode locations. All electrodes were referred to the nasion. EEG signals were recorded at 2,048Hz and down-sampled to 256 Hz for further analysis.

For each subject, six 4s-long segments with minimum artifacts were extracted manually from each 30-s block. The segments corresponding to the same stimulus frequency across four sessions were concatenated into a 96-s-long segment. Four occipital channels (F26, F28, F29, and F31) located around the mid-occipital location (Oz) were selected for SSVEP analysis, since SSVEPs measured over the visual cortex have the highest SNR [9]. Next, the 4-channel 96-s data segment was separated in turn into *N*-s epochs (*N*=1, 2, and 4 second (s)) for comparison.

C. CCA-based SSVEP frequency recognition

Lin *et al*. [4] first introduced CCA to detect the frequency of SSVEPs, which outperformed power spectral density (PSD)-based method. CCA theoretically aims to find a maximal correlation coefficient between two multivariate time series. That is, CCA identifies the SSVEP frequency by finding the maximal correlation ρ_f between multi-channel EEG signals and predefined sinusoidal reference signals associated with each flickering frequency.

The conventional CCA-based frequency detection can be expressed as follows:

$$\hat{f}_s = \arg\max_f \ \rho_f = \arg\max_f \ \rho(X, Y_f)$$

$$Y_f = \begin{bmatrix} \sin(2\pi ft) \\ \cos(2\pi ft) \end{bmatrix}$$

where X is the 4-channel SSVEP signal and Y_f is the reference signal corresponding to frequency f.

D. Differential Canonical Correlation Analysis

The efficiency of the conventional CCA in the detection of high-frequency SSVEP would presumably be confined by the power-law distribution of EEG spectra. It is not surprising to see that the canonical correlation tends to degrade as the flickering frequency increases, which was reported in a previous study [5]. To alleviate the interference from the spontaneous background EEG activities, this study proposes a new SSVEP detection method, named differential CCA (dCCA), to quantify the amount of canonical correlation purely elicited by target flickers. The basic idea is that if the acquired SSVEP signals are specifically filtered by a notch filter at the target frequency, the canonical correlation value calculated by CCA is supposed to drop significantly. That is, a largest drop of canonical correlation caused by the notch filtering appears at the target frequency. Accordingly, the proposed dCCA method aims to assess the differential canonical correlation values associated with each of the flickering frequencies, which could largely eliminate the impact of the power-law bias in EEG spectrum on SSVEP detectability.

First, 4-channel *N*-s SSVEP signals from each subject underwent filtering using IIR notch filters with a bandwidth of $1/W_n$ Hz centered at the stimulation frequencies, where W_n was a scale determining the sharpness of the notch filter in frequency response. Prior to notch filtering, baseline subtraction was applied to each channel of EEG data to remove the baseline drift and voltage offset estimated by 100-point moving average. To reduce the transient response of the notch filter, we empirically duplicated the *N*-s epoch four times and appended them into a $4 \times N$-s epoch. A notch filter with zero-phase shift was then applied to the appended epoch. The segment within [$2N$ $3N$] was extracted for CCA calculation. This process repeated five times at each flickering frequency to derive five notch-filtered epochs (X_{-f}). One constraint of the duplicate-padding method is the phase continuity between the duplicates, i.e. the number of stimulating periods in each duplicate has to be an integer so that SSVEPs in the appended duplicates have continuous phase throughout the whole epoch.

This study defined the canonical correlation between the filtered dataset X_{-f} and the reference signal Y_f as $\rho(X_{-f},Y_f)$ or ρ_{-f}. The ratio ρ_{-f}/ρ_f, i.e., differential canonical correlation, reflects the change of canonical correlation between the filtered dataset (turn-off) and the unfiltered dataset (turn-on). For simplicity, the ratio ρ_{-f}/ρ_f was named off-on-ratio (OOR) hereinafter. If OOR is small, there is a drop in canonical correlation as the power at f has been dramatically attenuated. In this way, OOR can be used as a new indicator to recognize the frequency of SSVEP:

$$\hat{f}_s = \arg\min_f \ OOR_f = \arg\min_f \ \frac{\rho_{-f}}{\rho_f}$$

Fig. 1 compares the effectiveness of using CCA and dCCA to identify a 1-s SSVEP epoch at 13Hz. This epoch was misclassified as 10 Hz when applying the conventional CCA method (max(ρ_f)). In contrast, the proposed dCCA method (min(OOR)) correctly identified the target frequency at 13Hz. That is, the drop of canonical correlation between ρ_{-f} and ρ_f was larger at 13Hz.

III. RESULTS

Figure 2 shows the averaged value of ρ, ρ_{-f}, and ρ_{-f}/ρ_f (OOR) for all trials across all subjects along different flickering frequencies. The results show that the canonical correlation peak, *i.e.* max(ρ), tended to monotonically decrease as flickering frequency increased (shown in Fig. 2(a), black line). In addition, as 13-Hz flickering was presenting, the 13-Hz peak among neighboring frequencies was not as distinguishable as the peaks induced by other lower flickering frequencies. The ρ_{-f} profile (gray line), derived by individually notch-filtering each frequency, exhibited a subtle descending slope toward higher frequency, which seemed to be identical to different stimuli. The above results lead to the gap between ρ and ρ_{-f} profiles capable of reflecting the onset frequency. The OOR based on the ratio of ρ_{-f} and ρ_f accordingly presented a

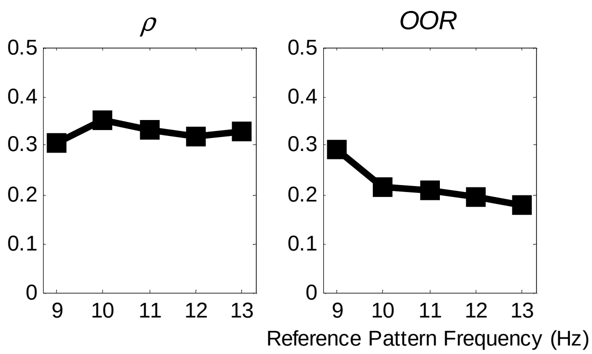

Fig. 1. The canonical correlation (ρ, left) and OOR (right) between an SSVEP epoch and different reference signals. The stimulating frequency of this epoch is 13Hz. The frequency was misrecognized as 10Hz when using max(ρ), while min(OOR) was able to recognize the stimulation frequency correctly.

distinct cave associated with the stimulating frequencies in Fig 2(b).

Table I summarizes the detection accuracy of 1-s SSVEP epochs using the CCA and dCCA approaches. The results show that dCCA returned better detectability than CCA in nine out of ten subjects. Across all subjects, the dCCA-based detectability was significantly higher than that of using CCA (67.85% vs. 61.52%, p<0.01, Wilcoxon paired sign-rank test). Tables II and III further summarize the confusion matrix of SSVEP detection performed by CCA and dCCA, respectively. For both methods, the SSVEP detectability along target frequency exhibited an explicit descending trend as frequency increased (see the diagonal elements of the confusion matrix). As expected, dCCA provided significantly improved detectability in higher frequency range (11 Hz: +6.77%; 12 Hz: +14.90%; 13 Hz: +21.88%). However, it also caused worse detectability at lower frequencies. As shown in Table IV, there was a noticeable drop at 9 Hz (-10.94%) and a marginal decrease at 10 Hz (-0.94%).

Fig. 3 presents the CCA- and dCCA-based classification results with different epoch lengths. Both methods obtained higher accuracy as the epoch length increased. Noticeably, dCCA consistently provided better performance than CCA across all conditions.

IV. Discussions and Conclusions

The conventional CCA-based SSVEP detection method is based on the assumption that the visual flickers evoke oscillations that have higher canonical correlation with the reference sinusoidal signal at the stimulating frequency. However, this might not robustly hold across different frequencies due to the power-law attenuation of SSVEP magnitude along stimulating frequency. The canonical correlation value ρ between SSVEPs and the reference signals decreased as flickering frequency increased (*c.f.* Fig. 2(a)), which was in line with previous work [5]. The higher frequencies (11-13 Hz) with low peak ρ were thus intensively mis-classified as lower frequencies (9 and 10 Hz) (c.f. Table II). In contrast, the proposed dCCA method was evidently capable of improving the discriminability of SSVEPs in higher frequency range (*c.f.* Fig. 2(b), Table III and Table IV), and thus provided statistically significantly improved performance than the conventional CCA-based method (*c.f.* Fig. 3). Unlike previous work based on machine-learning methods [7], the proposed dCCA method only uses a simple, training-free signal processing procedure (i.e., notch filtering), and thus can promote the feasibility and practicality of real-life applications of SSVEP BCIs.

The reason of using notch filtering was based on the assumption that eliminating the EEG signal at the stimulating frequency would cause a large drop in canonical correlation compared with other frequencies. It is worth noting that although the notch-filtering facilitates the detection of SSVEPs in high frequency range, transient response of notch filters would affect the effectiveness of dCCA to some extent, especially when using it on EEG epochs with a short duration. To solve this problem, this study alternatively duplicated the short epochs four times and concatenated them into a longer epoch. However, to keep the phase continuity in the filtered

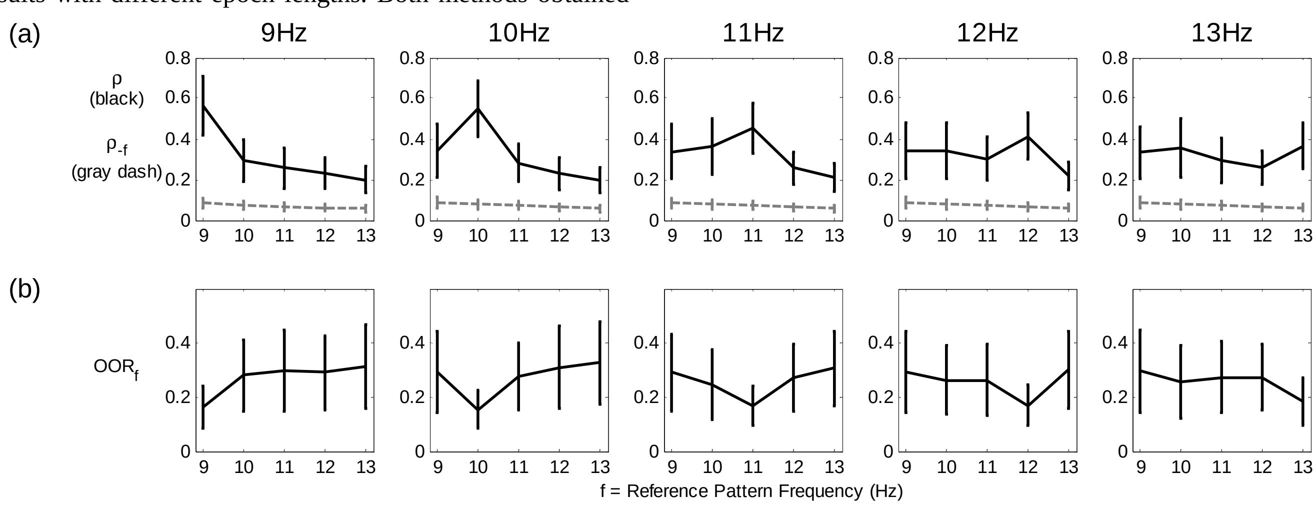

Fig. 2. (a) A comparison between ρ_f and ρ_{-f} given different stimulating frequencies (9-13 Hz) with 1s epoch length. Solid black line: the average ρ_f accross subjects; Dash gray line: the average ρ_{-f} accross subjects. Errorbars indicate the standard deviation. (b) The average OOR accross subjects. The stimulating frequency was presented at the top of each column.

TABLE I. THE ACCURACY OF SSVEP DETECTION USING CCA AND DCCA

Subject	Accuracy (%)	
	CCA - max(ρ)	dCCA - min(OOR)
S1	61.46	73.54
S2	55.00	57.50
S3	42.71	58.54
S4	58.75	67.29
S5	85.63	86.46
S7	51.25	59.58
S8	62.29	69.17
S9	72.50	72.08
S11	84.79	91.25
S12	40.83	43.13
Mean	61.52	67.85
SD	14.77	13.49

Epoch Length = 1s

TABLE II. CONFUSION MATRIX USING CCA

(%)		Estimated				
		9Hz	10Hz	11Hz	12Hz	13Hz
Target	9Hz	**88.54**	4.90	3.13	2.19	1.25
	10Hz	11.35	**82.60**	4.06	1.67	0.31
	11Hz	15.83	24.38	**54.90**	3.75	1.15
	12Hz	20.83	19.48	10.83	**47.40**	1.46
	13Hz	21.88	28.02	10.31	5.63	**34.17**

Total Accuracy: 61.52%; Epoch Length = 1s

TABLE III. CONFUSION MATRIX USING DCCA

(%)		Estimated				
		9Hz	10Hz	11Hz	12Hz	13Hz
Target	9Hz	**77.60**	6.88	6.04	5.31	4.17
	10Hz	6.98	**81.67**	6.04	3.23	2.08
	11Hz	9.38	16.88	**61.67**	7.29	4.79
	12Hz	9.90	14.06	9.90	**62.29**	3.85
	13Hz	8.54	18.23	9.69	7.50	**56.04**

Total Accuracy: 67.85; Epoch Length = 1s

TABLE IV. DIFFERENCE IN CONFUSION MATRIX BETWEEN USING (1)CCA AND (2) DCCA

(2) - (1) (%)		Estimated				
		9Hz	10Hz	11Hz	12Hz	13Hz
Target	9Hz	**-10.94**	1.98	2.92	3.13	2.92
	10Hz	-4.38	**-0.94**	1.98	1.56	1.77
	11Hz	-6.46	-7.50	**6.77**	3.54	3.65
	12Hz	-10.94	-5.42	-0.94	**14.90**	2.40
	13Hz	-13.33	-9.79	-0.63	1.88	**21.88**

Epoch Length = 1s

signal, this approach is subjected to the condition that the number of periods per epoch is an integer. As a result, we were unable to implement the dCCA method with 0.5-s epochs.

Currently, most of the SSVEP-based BCI systems adopted flickers with relatively low frequencies to elicit SSVEPs. This can be attributed to the fact that low-frequency SSVEPs usually have better SNR and detectability (e.g., the alpha frequency band). However, the perceivable visual flickers at low frequencies would inevitably cause more visual fatigue at the same time. The proposed dCCA method possesses the characteristics of alleviating the impact from the power-law attenuation of SSVEP SNR and is able to thereby increase the

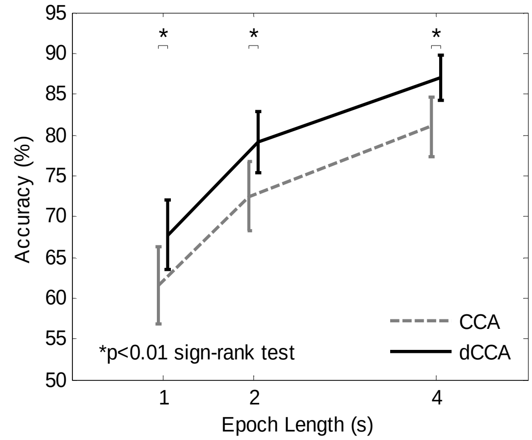

Fig. 3. The performance of SSVEP frequency recognition by using two strategies given epoch length = 1, 2, and 4 seconds. Error bars indicate standard error.

detectability of SSVEPs in relatively high frequencies. A natural next step is to assess the feasibility of applying dCCA to detect the SSVEPs in the frequency range higher than the human visual perception level (e.g., >30 Hz), which could considerably relieve the visual fatigue in operating SSVEP-based BCI systems.

REFERENCES

[1] Y.-T. Wang, Y. Wang, and T.-P. Jung, "A cell-phone-based brain–computer interface for communication in daily life," *J. Neural Eng.*, vol. 8, no. 2, p. 025018, Apr. 2011.

[2] 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.

[3] 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.

[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. 53, no. 12, pp. 2610–2614, 2006.

[5] 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 *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2012, pp. 1810–1814.

[6] Y. Wang, R. Wang, X. Gao, and S. Gao, "Brain-computer interface based on the high-frequency steady-state visual evoked potential," in *2005 First International Conference on Neural Interface and Control, 2005. Proceedings*, 2005, pp. 37–39.

[7] T. Tanaka, C. Zhang, and H. Higashi, "SSVEP frequency detection methods considering background EEG," in *2012 Joint 6th International Conference on Soft Computing and Intelligent Systems (SCIS) and 13th International Symposium on Advanced Intelligent Systems (ISIS)*, 2012, pp. 1138–1143.

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

[9] Y.-T. Wang, Y. Wang, C.-K. Cheng, and T.-P. Jung, "Measuring Steady-State Visual Evoked Potentials from non-hair-bearing areas," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2012, pp. 1806–1809.