

# Session-to-Session Transfer in Detecting Steady-State Visual Evoked Potentials with Individual Training Data

Masaki Nakanishi<sup>1</sup>(✉), Yijun Wang<sup>1,2</sup>, and Tzyy-Ping Jung<sup>1</sup>

<sup>1</sup> Swartz Center for Computational Neuroscience, Institute for Neural Computation, University of California San Diego, La Jolla, CA, USA  
{masaki, jung}@sccn.ucsd.edu

<sup>2</sup> State Key Laboratory on Integrated Optoelectronics, Institute of Semiconductors, Chinese Academy of Science, Beijing, China  
wangyj@semi.ac.cn

**Abstract.** The information transfer rate (ITR) of steady-state visual evoked potential (SSVEP)-based brain-computer interfaces (BCIs) has been significantly improved in the past few years. Recent studies have demonstrated the efficacy of advanced signal processing methods, which incorporate preliminarily recorded individual training data in SSVEP detection. However, conducting experiments for collecting training data from each individual is cumbersome because it is time-consuming and may cause visual fatigue. To simplify the training procedure, this study employs a session-to-session transfer method, which uses transfer templates obtained from datasets collected from the same subjects on a different day. The proposed approach was evaluated with a 40-class SSVEP dataset from eight subjects, each participated in two sessions on two different days. Study results showed that the proposed transfer method achieved significantly higher performance than conventional method based on canonical correlation analysis (CCA). In addition, by employing online adaptation, the proposed method reached high performance that is comparable with the most efficient approach in previous studies. These results indicate the feasibility of a high-performance SSVEP-based BCI with no or little training.

**Keywords:** Brain-computer interfaces (BCI) · Canonical correlation analysis (CCA) · Electroencephalogram (EEG) · Steady-state visual evoked potentials (SSVEP) · Transfer learning

## 1 Introduction

Steady-state visual evoked potentials (SSVEPs) are the brain's electrical responses to flickering visual stimuli. SSVEPs have been widely used in electroencephalogram (EEG)-based brain-computer interface (BCI) systems due to the advantage of high information transfer rate (ITR) [1, 2]. In SSVEP-based BCIs, users are required to gaze at one of multiple visual stimuli tagged with different stimulation properties such as frequencies and/or phases [1]. A target visual stimulus can be identified through analyzing the elicited SSVEPs using target identification methods. In this way, an SSVEP-based BCI can translate intentional brain activities into commands to control external devices.

Performance of SSVEP-based BCIs can be attributed to several factors including stimulus presentation, multiple target coding, and target identification algorithm [3]. In recent studies, advanced stimulus presentation and target coding methods significantly increased the number of stimuli that can be presented on a computer monitor [3–6]. For instance, 32-target and 40-target spellers that employed hybrid frequency and phase coding were designed [3, 6, 7]. The target identification method also plays an important role in improving the performance of SSVEP-based BCIs. Recent advances in signal processing methods, which incorporate individual training data, have been proposed to improve the performance of SSVEP detection [8]. Although these methods achieved better performance than conventional training-free methods, the training procedure can be time consuming and may cause visual fatigue because multiple trials have to be recorded before online operation.

To address this issue, Yuan et al. employed a transfer-learning approach, which transfers SSVEP templates from existing subjects to a new subject, and demonstrated the effectiveness of the method to improve the target identification accuracy compared with other training-free approaches [9]. However, since it is known that there is individual difference in anatomical shape and extent of area V1, where the source of SSVEPs is located at [10, 11], the performance improvement of transferring SSVEP data from different subjects might be limited. Therefore, this study employs a session-to-session transfer method (i.e., training data collected from the same subjects on a different day) to reduce training time and investigated the performance of the proposed approach in terms of the classification accuracy and ITRs.

## 2 Material and Method

### 2.1 Experimental Design

EEG data were recorded in a simulated online BCI experiment. 40 visual stimuli were presented on a 23.6-in. liquid-crystal display screen with a resolution of  $1,920 \times 1,080$  pixels and a refresh rate of 60 Hz. The stimuli were arranged in a  $5 \times 8$  matrix, and tagged with 40 different frequencies (8.0 Hz to 15.8 Hz with an interval of 0.2 Hz) with 4 different phases (0,  $0.5\pi$ ,  $\pi$ ,  $1.5\pi$ ). The stimulation frequencies and phases were generated using the joint frequency-phase modulation (JFPM) method [7].

Eight healthy subjects with normal or corrected-to-normal vision participated in this study. Each subject performed the experiment on two different days. All subjects read and signed an informed consent form before participating in the experiment. The subjects sat in a comfortable chair positioned approximately 70 cm from a computer monitor, and gazed at one of the visual stimuli. From each subject, six 5 s-long data and fifteen 1 s-long data of SSVEPs for each visual stimulus were recorded in two experiments conducted on different days. The intervals of two experiment days differed across individuals. After stimulus offset, the screen was blank for 0.5 s before the next trial began. EEG data were acquired using a Synamps2 system (Neuroscan, Inc.) at a sampling rate of 1,000 Hz. Nine electrodes placed over the parietal and occipital areas (Pz, PO5, PO3, POz, PO4, PO6, O1, Oz, and O2) were used to measure SSVEPs.

## 2.2 Target Identification

This study employed the canonical correlation analysis (CCA) based target identification algorithm with individual calibration templates and transferred templates from a different day [8, 9]. The online transferred template-based CCA proposed by Yuan et al., which updates the transferred templates adaptively in online operation, was also tested in this study [9]. In addition, we tested the standard CCA, which is an unsupervised approach, as a comparative method. In all methods, filter bank analysis was applied [12].

**Standard CCA.** CCA has been widely used to detect the frequency of SSVEPs [13, 14]. In the CCA-based SSVEP detection method, canonical correlation between multichannel EEG signals  $\mathbf{X} \in \mathbb{R}^{N_c \times N_s}$  and sine-cosine reference signals  $\mathbf{Y} \in \mathbb{R}^{2N_h \times N_s}$  are calculated as:

$$\rho(\mathbf{X}, \mathbf{Y}) = \max_{\mathbf{w}_x, \mathbf{w}_y} \frac{E[\mathbf{w}_x^T \mathbf{X} \mathbf{Y}^T \mathbf{w}_y]}{\sqrt{E[\mathbf{w}_x^T \mathbf{X} \mathbf{X}^T \mathbf{w}_x] \cdot E[\mathbf{w}_y^T \mathbf{Y} \mathbf{Y}^T \mathbf{w}_y]}} \quad (1)$$

Here,  $N_c$ ,  $N_s$  and  $N_h$  denote the number of channels, the number of sampling points, and the number of harmonics being considered, respectively. The maximum of  $\rho$  with respect to  $\mathbf{w}_x$  and  $\mathbf{w}_y$  is the maximum canonical correlation. The reference signal  $\mathbf{Y}_{f_n}$  corresponding to the stimulus frequency  $f_n$  are defined as:

$$\mathbf{Y}_{f_n} = \begin{bmatrix} \sin(2\pi f_n t) \\ \cos(2\pi f_n t) \\ \vdots \\ \sin(2\pi N_h f_n t) \\ \cos(2\pi N_h f_n t) \end{bmatrix}, t = [1, 2, \dots, N_s] \cdot \frac{1}{f_s}. \quad (2)$$

Where,  $f_s$  is the sampling frequency. The frequency of the reference signals with maximal correlation was selected as the frequency of the SSVEPs.

**CCA with Individual Training Data.** Our recent studies proposed an extended CCA-based method which incorporates individual training data [3, 8, 15]. This method exploits important signal characteristics from existing training data for improving the target identification. In this method, a spatial filter  $\mathbf{w}_{xy}$  that maximizes the signal-to-noise ratio (SNR) of training set  $\hat{\mathbf{X}}_n$  for  $n$ -th target was first obtained by performing CCA with  $\hat{\mathbf{X}}_n$  and  $\mathbf{Y}_{f_n}$ . Also, a spatial filter  $\mathbf{w}_x$  that maximizes the SNR of test EEG data  $\mathbf{X}$  was obtained by performing CCA with  $\mathbf{X}$  and  $\mathbf{Y}_{f_n}$ . After that, Pearson's correlation coefficients between test data  $\mathbf{X}$  and training data  $\hat{\mathbf{X}}_n$  projected onto these two spatial filters and CCA with test data  $\mathbf{X}$  and reference signal  $\mathbf{Y}_{f_n}$  were calculated as:

$$\mathbf{r}_n = \begin{bmatrix} r_{n,1} \\ r_{n,2} \\ r_{n,3} \end{bmatrix} = \begin{bmatrix} r(\mathbf{X}^T \mathbf{w}_{xy}, \widehat{\mathbf{X}}_n^T \mathbf{w}_{xy}) \\ r(\mathbf{X}^T \mathbf{w}_{xy}, \widehat{\mathbf{X}}_n^T \mathbf{w}_{xy}) \\ r(\mathbf{X}^T \mathbf{w}_x, \mathbf{Y}_n^T \mathbf{w}_y) \end{bmatrix}. \quad (3)$$

Where,  $r(a, b)$  indicates the Pearson's correlation analysis between two one-dimensional signals  $a$  and  $b$ . The following ensemble classifier can be used as the final feature in target identification:

$$\rho_n = \sum_{i=1}^3 \text{sign}(r_{n,i}) \cdot r_{n,i}^2. \quad (4)$$

The template  $\widehat{\mathbf{X}}_n$  that maximizes the weighted correlation value  $\rho_n$  is selected as the reference signal corresponding to the target. In this study, two training datasets were used to make separate templates to evaluate the intra-day and inter-day variability in the SSVEP detection. These methods are termed individual template-based CCA (it-CCA) and transfer template-based CCA (tt-CCA), respectively [9].

**Online transfer-template CCA (ott-CCA).** In this study, the online tt-CCA (ott-CCA) proposed by Yuan et al. was also tested to investigate the efficacy of online adaptation in session-to-session transfer SSVEP detection [9]. The ott-CCA updates the transferred templates in online operation as follows: (1) Calculate the difference between the first and second largest feature value  $\rho_n$  among all candidate targets, (2) If the difference is higher than a pre-defined threshold  $thr$ , update the template via Eq. (5).

$$\widehat{\mathbf{X}}_n^{\text{new}} = \frac{1}{m_n^{\text{new}}} \left( m_n^{\text{old}} \widehat{\mathbf{X}}_n^{\text{old}} + \mathbf{X} \right) \quad (5)$$

where  $\widehat{\mathbf{X}}_n^{\text{old}}$  and  $\widehat{\mathbf{X}}_n^{\text{new}}$  are old and new templates for  $n$ -th target.  $m_n^{\text{old}}$  is the number of the trials that have been used to get the averaged template and  $m_n^{\text{new}} = m_n^{\text{old}} + 1$ . According to the previous study [9], the threshold  $thr$  was set to 0.1.

**Performance Evaluation.** In this study, the performance of aforementioned methods was tested with the test dataset, which consisted of nine (the seventh to fifteenth) trials from day 2. To evaluate the performance of session-to-session transfer learning approaches, we prepared two separate templates from first six trials from day 2 for it-CCA, and all six trials from day 1 for tt-CCA.

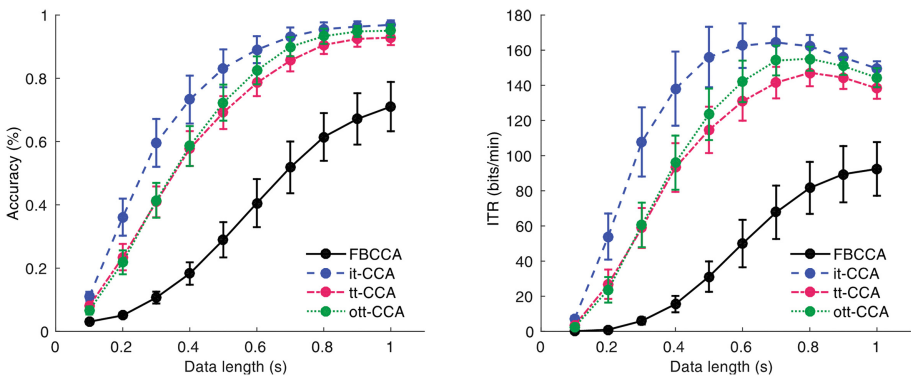
The performance was evaluated by the target identification accuracy and information transfer rate (ITR) calculated as [16]:

$$ITR = \left( \log_2 N_f + P \log_2 P + (1 - P) \log_2 \left[ \frac{1 - P}{N_f - 1} \right] \right) \times \left( \frac{60}{T} \right) \quad (6)$$

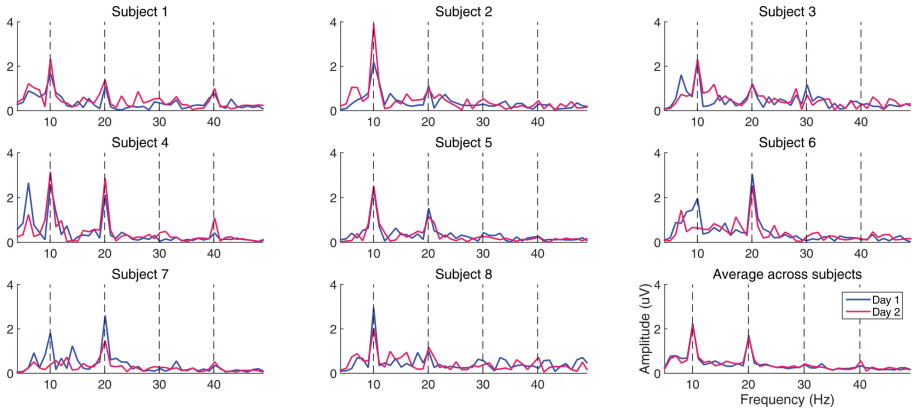
where  $N_f$  is the number of visual stimulus (i.e.,  $N_f = 40$  in this study),  $P$  is the target identification accuracy, and  $T$  (seconds/selection) is the average time for a selection. This study calculated the performance with different  $T$  (Target gazing time: 0.1 s to 1.0 s with an interval of 0.1 s; Gaze shifting time: 1.0 s).

### 3 Results

Figure 1 shows the averaged accuracy of target identification and the ITR across subjects with different data lengths. In general, template-based CCA methods outperformed the filter bank CCA (FBCCA) regardless of the data length. One-way repeated measure analysis of variance (ANOVA) showed there was significant difference in the target identification accuracy between four methods under all data length ( $p < 0.05$ ). Although the performance of tt-CCA was lower than that of it-CCA, tt-CCA significantly improved the performance over FBCCA. With longer data length (e.g.,  $>0.6$  s), ott-CCA significantly improved the accuracy over tt-CCA (ott-CCA vs. tt-CCA; 0.6 s:  $82.50 \pm 4.42\%$  vs.  $78.54 \pm 4.20\%$ ,  $p < 0.05$ , 0.7 s:  $89.97 \pm 3.02\%$  vs.  $85.59 \pm 3.42\%$ ,  $p < 0.05$ , 0.8 s:  $93.30 \pm 2.42\%$  vs.  $90.56 \pm 2.89\%$ ,  $p < 0.05$ , 0.9 s:  $94.79 \pm 1.77\%$  vs.  $92.43 \pm 2.50\%$ ,  $p = 0.05$ , 1.0 s:  $95.03 \pm 2.02\%$  vs.  $92.85 \pm 2.35\%$ ,  $p < 0.05$ ). Although there was significant difference in the accuracy between ott-CCA and it-CCA, ott-CCA achieved comparable accuracy to it-CCA (ott-CCA vs. it-CCA; 0.9 s:  $94.79 \pm 1.77\%$  vs.  $96.32 \pm 1.67\%$ ,  $p < 0.05$ , 1.0 s:  $95.03 \pm 2.02\%$  vs.  $96.88 \pm 1.41\%$ ,  $p < 0.05$ ). The difference of ITRs between these methods was consistent with that of the accuracy. The data length corresponding to the highest ITR was different for each method (FBCCA: 1.0 s,  $92.451 \pm 5.26$  bits/min, it-CCA: 0.7 s,  $164.38 \pm 8.93$  bits/min, tt-CCA: 0.8 s,  $147.16 \pm 7.68$  bits/min, ott-CCA: 0.8 s,  $155.10 \pm 6.84$  bits/min).



**Fig. 1.** Averaged accuracy of target identification and ITRs across subjects with different data lengths. The error bars indicate standard errors.



**Fig. 2.** Amplitude spectra of SSVEP signals at 10 Hz recorded from different days for each subject. The dash lines indicate the fundamental and harmonic frequencies (Color figure online).

## 4 Discussions

To compare the difference of SSVEP characteristics recorded from different days, the amplitude spectra of SSVEPs were calculated by the fast Fourier transform (FFT). Figure 2 depicts examples of the amplitude spectra of single-channel SSVEPs at 10 Hz for all eight subjects and the average spectrum. The spectra show that the fundamental and harmonics frequency components have higher amplitude than the background EEG signals. Interestingly, in two out of eight subjects (Subjects 6 and 7), the amplitude at the second harmonic frequency was higher than that at the fundamental frequency on both days. Despite of this consistency in the fundamental and harmonic components in response to flickering visual stimuli, background EEG signals are different between different days. Therefore, to improve the performance of session-to-session transfer learning in SSVEP-based BCIs, background signals should be removed prior to CCA. Although it might be nearly impossible to put electrodes on exactly the same location between different days, the CCA-based spatial filtering could enhance the performance of template matching by extracting components being correlated with artificially generated reference signals.

The BCI performance in the present study showed that tt-CCA can achieve significantly higher target identification accuracy and ITR than FBCCA. Although the performance of tt-CCA was lower than it-CCA, it can be improved to a comparable level as it-CCA by employing the online adaptation (ott-CCA) with long data length. These results indicate that collecting training data was no longer required to optimize the performance of SSVEP-based BCIs since after the first run. By combining subject-to-subject transfer template [9] with the session-to-session transfer learning, higher BCI performance could be obtained without any preliminary experiment to record training data.

## 5 Conclusion

This study showed that a session-to-session transfer method could facilitate the training procedure in an SSVEP-based BCI using individual calibration data. In addition, the adaptive approach can further optimize individual templates while users are operating the system. These findings suggest that session-to-session transfer is efficient for implementing a high-speed SSVEP-based BCI system with zero training.

**Acknowledgement.** This research was supported in part by a gift fund from Swartz Foundation, by Army Research Laboratory (W911NF-10-2-0022), DARPA (US-DID11PC20183), and UC Proof of Concept Grant Award (269228).

## References

1. Wang, Y., Gao, X., Hong, B., Jia, C., Gao, S.: Brain-computer interfaces based on visual evoked potentials: feasibility of practical system design. *IEEE Eng. Med. Biol. Mag.* **27**(5), 64–71 (2008)
2. Gao, S., Wang, Y., Gao, S., Hong, B.: Visual and auditory brain-computer interfaces. *IEEE Trans. Biomed. Eng.* **61**(5), 1436–1447 (2014)
3. Nakanishi, M., Wang, Y., Wang, Y.T., Mitsukura, Y., Jung, T.P.: A high-speed brain speller using steady-state visual evoked potentials. *Int. J. Neural Syst.* **24**(6), 1450019 (2014)
4. Wang, Y., Wang, Y.T., Jung, T.P.: Visual stimulus design for high-rate SSVEP. *Electron. Lett.* **46**(15), 1057–1058 (2010)
5. Nakanishi, M., Wang, Y., Wang, Y.T., Mitsukura, Y., Jung, T.P.: Generating visual flickers for eliciting robust steady-state visual evoked potentials at flexible frequencies using monitor refresh rate. *PLoS ONE* **9**(6), e99235 (2014)
6. Chen, X., Chen, Z., Gao, S., Gao, X.: A high-ITR SSVEP based BCI speller. *Brain-Comp. Interfaces* **1**(3–4), 181–191 (2014)
7. Chen, X., Wang, Y., Nakanishi, M., Gao, X., Jung, T.P., Gao, S.: High-speed spelling with a noninvasive brain-computer interface. *Proc. Natl. Acad. Sci. U. S. A.* **112**(44), E6058–E6067 (2015)
8. Nakanishi, M., Wang, Y., Wang, Y.T., Jung, T.P.: A comparison study of canonical correlation analysis based methods for detecting steady-state visual evoked potentials. *PLoS ONE* **10**(10), e140703 (2015)
9. Yuan, P., Chen, X., Wang, Y., Gao, X., Gao, S.: Enhancing performances of SSVEP-based brain-computer interfaces via exploiting inter-subject information. *J. Neural Eng.* **12**, 046006 (2015)
10. Brindley, G.: The variability of the human striate cortex. *J. Physiol.* **225**, 1–3 (1972)
11. Stensaas, S., Eddington, D., Dobbelle, W.: The topography and variability of the primary visual cortex in man. *J. Neurosurg.* **40**, 747–755 (1974)
12. Chen, X., Wang, Y., Gao, S., Jung, T.P., Gao, X.: Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain-computer interface. *J. Neural Eng.* **12**, 046008 (2015)
13. Lin, Z., Zhang, C., Wu, W., Gao, X.: Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs. *IEEE Trans. Biomed. Eng.* **52**, 1172–1176 (2007)

14. Bin, G., Gao, X., Yan, Z., Hong, B., Goa, S.: An online multi-channel SSVEP-based brain-computer interface using a canonical correlation analysis method. *J. Neural Eng.* **6**, 046002 (2009)
15. Wang, Y., Nakanishi, M., Wang, Y.T., Jung, T.P.: Enhancing detection of steady-state visual evoked potentials using individual training data. In: 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 3037–3040. Chicago, IL (2014)
16. Cheng, M., Gao, X., Gao, S., Xu, D.: Design and implementation of a brain-computer interface with high transfer rate. *IEEE Trans. Biomed. Eng.* **49**(10), 1181–1186 (2002)