L1-Regularized Multiway Canonical Correlation Analysis for SSVEP-Based BCI

Yu Zhang, *Member, IEEE*, Guoxu Zhou, *Member, IEEE*, Jing Jin, Minjue Wang, Xingyu Wang, and Andrzej Cichocki, *Fellow, IEEE*

***Abstract*—Canonical correlation analysis (CCA) between recorded electroencephalogram (EEG) and designed reference signals of sine-cosine waves usually works well for steady-state visual evoked potential (SSVEP) recognition in brain–computer interface (BCI) application. However, using the reference signals of sine-cosine waves without subject-specific and inter-trial information can hardly give the optimal recognition accuracy, due to possible overfitting, especially within a short time window length. This paper introduces an L1-regularized multiway canonical correlation analysis (L1-MCCA) for reference signal optimization to improve the SSVEP recognition performance further. A multiway extension of the CCA, called MCCA, is first presented, in which collaborative CCAs are exploited to optimize the reference signals in correlation analysis for SSVEP recognition alternatingly from the channel-way and trial-way arrays of constructed EEG tensor. L1-regularization is subsequently imposed on the trial-way array optimization in the MCCA, and hence results in the more powerful L1-MCCA with function of effective trial selection. Both the proposed MCCA and L1-MCCA methods are validated for SSVEP recognition with EEG data from 10 healthy subjects, and compared to the ordinary CCA without reference signal optimization. Experimental results show that the MCCA significantly outperforms the CCA for SSVEP recognition. The L1-MCCA further improves the recognition accuracy which is significantly higher than that of the MCCA.**

***Index Terms*—Brain–computer interface (BCI), electroencephalogram (EEG), l1-regularization, multiway canonical correlation analysis (MCCA), steady-state visual evoked potential (SSVEP).**

I. Introduction

A BRAIN–COMPUTER interface (BCI) is a communication system that allows a direct connection between the human brain and computer, and hence provides a communication channel for people with severe motor disabilities to reestablish environmental control abilities [1]. In recent years, BCI has been developed mainly based on three types of brain activities measured by electroencephalogram (EEG), i.e., event-related potential (ERP), sensorimotor rhythm (SMR), and steady-state visual evoked potential (SSVEP) [2]–[6]. Since the SSVEP-based BCI usually provides relatively higher information transfer rate (ITR) and requires less training to the user, it has been increasingly studied [7], [8].

SSVEP is evoked over occipital scalp region with the same frequency as that of the visual stimulus and may also include its harmonics when subject focuses on the repetitive flicker of a visual stimulus [9]. With this characteristic, the SSVEP-based BCI is developed to detect commands desired by the subject through recognizing the SSVEP frequency components in EEG signals. Since the SSVEP is likely to be contaminated by spontaneous EEG and other background noises in the brain, how to recognize its frequency components with a high accuracy, especially at a short time window length (TW), is a challenge issue and considerably important for development of an improved SSVEP-based BCI.

Power spectral density analysis (PSDA) is a traditional method for SSVEP recognition, which estimates the PSD values of EEG signals at different frequencies within a specific TW, typically by fast Fourier transform (FFT) [10]. The frequency with the maximal PSD value is recognized as the target frequency (i.e., the SSVEP frequency component). Two problems of the PSDA for SSVEP recognition are: 1) PSDA with a single or bipolar channel may be sensitive to noises, and hence obtain a low signal-to-noise ratio (SNR) [11], [12]; 2) a relatively long TW (e.g., $> 3s$) is usually required to estimate the spectrum with sufficient frequency resolution, which therefore limits the real-time performance of SSVEP-based BCI to some extent [13], [14]. To overcome drawbacks of the PSDA, several advance approaches have been proposed to improve the SSVEP recognition performance [11]–[13], [15], in which a canonical correlation analysis (CCA) based recognition method, first introduced by Lin *et al.* [11], has aroused more interests of researchers. With the CCA, correlation was maximized between the EEG signals from multiple channels and the reference signals of sine-cosine at each of the used stimulus frequencies. The stimulus frequency with the maximal correlation coefficient was recognized as the target frequency. The use of CCA provided better SSVEP recognition performance than that of the PSDA since it delivered an optimization for the combination of multiple channels to improve the SNR. Recently, the CCA has been widely adopted in the developments of SSVEP-based BCIs [16]–[18] and other related BCIs [19]–[21].

Manuscript received April 21, 2013; revised July 23, 2013; accepted August 21, 2013. Date of publication October 07, 2013; date of current version November 04, 2013. This work was supported in part by the Nation Nature Science Foundation of China under Grant 61305028, Grant 61074113, Grant 61203127, and Grant 61103122, in part by the Fundamental Research Funds for the Central Universities under Grant WH1314023 and Grant WH1114038, and in part Shanghai Leading Academic Discipline Project B504.

Y. Zhang, J. Jin, M. Wang, and X. Wang are with the Key Laboratory for Advanced Control and Optimization for Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China (e-mail: zhangyu0112@gmail.com; jinjingat@gmail.com; wangminjue@163.com; xywang@ecust.edu.cn).

G. Zhou is with the Laboratory for Advanced Brain Signal Processing, RIKEN Brain Science Institute, Wako-shi, Saitama 351-0198, Japan (e-mail: zhouguoxu@brain.riken.jp).

A. Cichocki is with the Laboratory for Advanced Brain Signal Processing, RIKEN Brain Science Institute, Wako-shi, Saitama 351-0198, Japan, and also the System Research Institute, Polish Academy of Sciences, Warsaw 00-901, Poland (e-mail: a.cichocki@riken.jp).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNSRE.2013.2279680

1534-4320 © 2013 IEEE

Although the CCA usually works well in SSVEP-based BCIs, the recognition performance may be deteriorated due to possible overfitting, especially within a short TW, when directly using the reference signals of sine-cosine waves without subject-specific and inter-trial information. Therefore, a sophisticated calibration processing are necessary to refine the reference signals from a certain number of trials for each specific subject, in order to further improve the SSVEP recognition performance. To this end, we introduce an L1-regularized multiway canonical correlation analysis (L1-MCCA) based on an ingenious combination of the tensor analysis [22]–[25] and the sparse regularization [26]–[28] to solve this problem. An unpenalized MCCA (conference version [14]) is first presented, in which collaborative CCAs are exploited to learn the projection vectors for reference signal optimization alternatingly from the channel-way and trial-way arrays of EEG tensor. In the MCCA, the EEG tensor is constructed by multi-channel EEG from multiple recording trials where some trials may bring more obstruction than contribution to the reference signal optimization. Thus, L1-regularization is further imposed on trial-way array optimization of the MCCA, which results in the more powerful L1-MCCA to learn sparse projection vector with function of effective trial selection during the reference signal optimization. Experimental results based on the EEG data from 10 subjects demonstrate that SSVEP recognition accuracy is significantly improved by the MCCA and further by the L1-MCCA for reference signal optimization, in compared to direct recognition method based on the CCA without refining for reference signals.

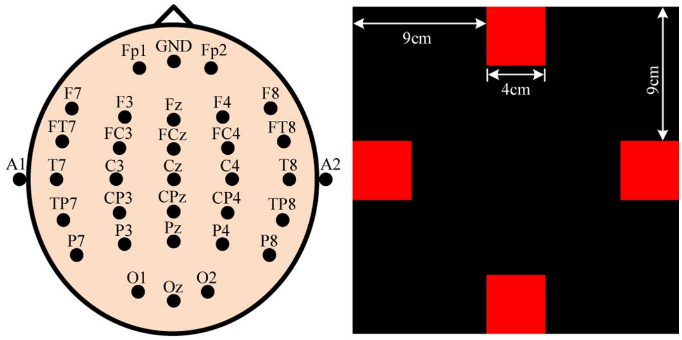
II. Materials and Methods

A. EEG Acquisition

1) Subjects: Ten healthy subjects (S1-S10, aged from 21 to 27 years, all males) participated in our experiment. All of them had normal or corrected to normal vision. The subjects were seated in a comfortable chair 60 cm from a standard 17 in CRT monitor (85 Hz refresh rate, 1024 × 768 screen resolution) in a shielded room.

2) EEG Recordings: EEG signals were recorded at 250 Hz sampling rate using the Nuamps amplifier (NuAmp, Neuroscan, Inc.) with high-pass and low-pass filters of 0.1 and 70 Hz from 30 channels that were placed on the standard positions according to the 10–20 international system [see Fig. 1(a)]. The average of two mastoid electrodes (A1, A2) was used as reference and the ground electrode (GND) was placed on the forehead. Since SSVEP is a localized potential and the occipital and parietal scalp areas have been demonstrated to contribute most to its recognition [11], [16], only the eight channels P7, P3, Pz, P4, P8, O1, Oz, and O2 are used for analysis in this study.

3) Experimental Paradigm: In the experiment, four red squares were flickered as stimuli at different four frequencies: 6, 8, 9, 10 Hz, respectively, on the black screen [see Fig. 1(b)]. Each subject completed 20 runs with 5–10 min break after the first 10 runs. The time interval between runs was 10 s. In each run, the subject was asked to focus on each of the four red squares once for 4 s, respectively, preceded by each target



(a) (b)

Fig. 1. Channel configuration for EEG recording (a) and SSVEP experimental layout (b).

cue duration (i.e., time interval between trials) of 2 s. A total of 80 trials (four trials in each run) EEG data were recorded from each subject. The EEG data were band-pass filtered from 4 to 45 Hz by a sixth-order forward–backward Butterworth bandpass filter.

B. CCA For SSVEP Recognition

Canonical correlation analysis (CCA) is a multivariable statistical method first proposed by Hotelling [29], which aims to reveal the underlying correlation between two sets of data. Consider two sets of random variables $\mathbf{X} \in \mathbb{R}^{I_1 \times J}$ and $\mathbf{Y} \in \mathbb{R}^{I_2 \times J}$, the CCA tries to find a pair of linear transforms $\mathbf{w} \in \mathbb{R}^{I_1}$ and $\mathbf{v} \in \mathbb{R}^{I_2}$ to maximize the correlation between linear combinations $\tilde{\mathbf{x}} = \mathbf{w}^T\mathbf{X}$ and $\tilde{\mathbf{y}} = \mathbf{v}^T\mathbf{Y}$, through solving the following optimization problem:

$$\begin{aligned}\rho &= \max_{\mathbf{w},\mathbf{v}} \frac{E[\tilde{\mathbf{x}}\tilde{\mathbf{y}}^T]}{\sqrt{E[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T]E[\tilde{\mathbf{y}}\tilde{\mathbf{y}}^T]}} \\ &= \max_{\mathbf{w},\mathbf{v}} \frac{\mathbf{w}^T\mathbf{X}\mathbf{Y}^T\mathbf{v}}{\sqrt{\mathbf{w}^T\mathbf{X}\mathbf{X}^T\mathbf{w}\mathbf{v}^T\mathbf{Y}\mathbf{Y}^T\mathbf{v}}} \end{aligned} \tag{1}$$

where ρ denotes the maximum correlation coefficient. The optimization problem in (1) can be solved by a generalized eigenvalue problem [30].

Lin *et al.* [11] first introduced the CCA into SSVEP recognition in BCI application. Assume there are M stimulus frequencies to be recognized in an SSVEP-based BCI. $\mathbf{X}$ denotes the EEG data recorded from I_1 channels with J points in each channel. $\mathbf{Y}_m$, as the reference signal at mth stimulus frequency f_m $(m = 1, 2, \ldots, M)$, is constructed by a series of sine-cosine waves as

$$\mathbf{Y}_m = \begin{pmatrix} \sin(2\pi f_m t) \\ \cos(2\pi f_m t) \\ \vdots \\ \sin(2\pi H f_m t) \\ \cos(2\pi H f_m t) \end{pmatrix}, \quad t = \frac{1}{F}, \frac{2}{F}, \ldots, \frac{J}{F} \tag{2}$$

where H denotes the number of used harmonics (i.e., $I_2 = 2H$) and F is the sampling rate. The second and third harmonics have been reported to be insignificant for SSVEP recognition in using the CCA method [16], and hence only the fundamental frequency component is considered in this study (i.e., $H = 1$).

Solving correlation coefficients between the EEG data and each of the reference signal by (1), the SSVEP target frequency f_t is recognized as

$$f_t = \max_{f_m} \rho_m, \quad m = 1, 2, \ldots, M. \tag{3}$$

C. MCCA for SSVEP Recognition

Although the CCA usually works quite well, the direct use of sine-cosine waves as reference signals in the correlation analysis for SSVEP recognition may not result in the optimal accuracy, due to their lack of subject-specific and inter-trial information. To improve the SSVEP recognition accuracy further, this study (conference version [14]) first introduces a multiway extension of the CCA, called MCCA, to optimize the reference signals through collaboratively maximizing correlation between the multiple dimensions (i.e., channel-way and trial-way arrays) of constructed EEG tensor and the sine-cosine waves.

To describe the MCCA and subsequent L1-regularized version conveniently, we first make a brief review for some preliminaries about tensor, i.e., the multiway array representation of data. The way of tensor is the number of dimensions [22], [31]. A one-way tensor is a vector and a two-way tensor is a matrix. An N-way tensor is denoted by $\mathcal{X} = (\mathcal{X})_{i_1 i_2 \ldots i_N} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$. The nth way projection of the tensor with a vector $\mathbf{w} \in \mathbb{R}^{I_n}$ is

$$\left(\mathcal{X} \times_n \mathbf{w}^T\right)_{i_1 \ldots i_{n-1} i_{n+1} \ldots i_N} = \sum_{i_n=1}^{I_n} x_{i_1 i_2 \ldots i_N} w_{i_n}. \tag{4}$$

Consider a three-way tensor $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ (channel $\times$ time $\times$ trial) constructed by multi-channel EEG data from multiple trials with a specific stimulus frequency and an original reference signal set $\mathbf{Y} \in \mathbb{R}^{2H \times J}$ constructed as (2). The MCCA tries to find linear transforms $\mathbf{w}_1 \in \mathbb{R}^I$, $\mathbf{w}_3 \in \mathbb{R}^K$, and $\mathbf{v} \in \mathbb{R}^{2H}$ to maximize the correlation between linear combinations $\tilde{\mathbf{x}} = \mathcal{X} \times_1 \mathbf{w}_1^T \times_3 \mathbf{w}_3^T$, and $\tilde{\mathbf{y}} = \mathbf{v}^T \mathbf{Y}$ as

$$\rho = \max_{\mathbf{w}_1, \mathbf{w}_3, \mathbf{v}} \frac{E[\tilde{\mathbf{x}} \tilde{\mathbf{y}}^T]}{\sqrt{E[\tilde{\mathbf{x}} \tilde{\mathbf{x}}^T] E[\tilde{\mathbf{y}} \tilde{\mathbf{y}}^T]}}. \tag{5}$$

An alternating algorithm based on the ordinary CCA is proposed to solve the optimization problem in (5), in which $\mathbf{w}_3$ is first fixed to solve $\mathbf{w}_1$ and $\mathbf{v}$, then $\mathbf{w}_1$ and $\mathbf{v}$ are fixed to solve $\mathbf{w}_3$ alternatingly, and repeats this procedure till a convergence criterion is satisfied. After the optimal linear transforms $\tilde{\mathbf{w}}_1$ and $\tilde{\mathbf{w}}_3$ are obtained, the optimized reference signal $\mathbf{z} \in \mathbb{R}^J$ is given by

$$\mathbf{z} = \mathcal{X} \times_1 \tilde{\mathbf{w}}_1^T \times_3 \tilde{\mathbf{w}}_3^T. \tag{6}$$

The proposed MCCA algorithm for reference signal optimization in SSVEP recognition is summarized in Algorithm 1.

With the optimized reference signal $\mathbf{z}_m$ corresponding to the stimulus frequency f_m ($m = 1, 2, \ldots, M$), the correlation between a new test data of single trial $\hat{\mathbf{X}} \in \mathbb{R}^{I \times J}$ and each of the reference signals is maximized by the ordinary CCA in (1), and the SSVEP target frequency is then recognized according to (3).

Algorithm 1: MCCA algorithm for reference signal optimization in SSVEP recognition

Input: EEG tensor $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ (channel $\times$ time $\times$ trial) recorded at a specific stimulus frequency and sine-cosine signals $\mathbf{Y} \in \mathbb{R}^{2H \times J}$ (harmonic $\times$ time) constructed as Eq. (2).
Output: Optimized reference signal $\mathbf{z} \in \mathbb{R}^J$

Random initialization for $\mathbf{w}_3 \in \mathbb{R}^K$;
Do $\tilde{\mathbf{X}} = \mathcal{X} \times_3 \mathbf{w}_3^T$;
repeat
  Solve $\mathbf{w}_1$, $\mathbf{v}$ by the CCA between $\tilde{\mathbf{X}}$ and $\mathbf{Y}$;
  $\mathbf{w}_1 = \mathbf{w}_1 / \|\mathbf{w}_1\|_2$, $\mathbf{v} = \mathbf{v} / \|\mathbf{v}\|_2$;
  $\tilde{\mathbf{X}} = \mathcal{X} \times_1 \mathbf{w}_1^T$, $\tilde{\mathbf{y}} = \mathbf{v}^T \mathbf{Y}$;
  Solve $\mathbf{w}_3$ by the CCA between $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{y}}$;
  $\mathbf{w}_3 = \mathbf{w}_3 / \|\mathbf{w}_3\|_2$;
  $\tilde{\mathbf{X}} = \mathcal{X} \times_3 \mathbf{w}_3^T$;
until *Stop criterion is met*;
$\tilde{\mathbf{w}}_1 = \mathbf{w}_1$, $\tilde{\mathbf{w}}_3 = \mathbf{w}_3$;
$\mathbf{z} = \mathcal{X} \times_1 \tilde{\mathbf{w}}_1^T \times_3 \tilde{\mathbf{w}}_3^T$.

D. L1-MCCA For SSVEP Recognition

In the MCCA, the projection vector of each dimension optimization is learned without regularization, and hence lack of sparsity that can provide greater interpretability for features [32]. In this section, a penalized MCCA with L1-regularization (L1-MCCA) is further proposed to learn sparse projection vector with function of automatic feature selection for reference signal optimization in SSVEP recognition.

Since any rescaling for $\mathbf{w}$ and $\mathbf{v}$ does not affect the correlation maximization formulated in (1), which is therefore equivalent to maximize $\mathbf{w}^T \mathbf{X} \mathbf{Y}^T \mathbf{w}$ (or minimize $-\mathbf{w}^T \mathbf{X} \mathbf{Y}^T \mathbf{w}$) subject to $\mathbf{w}^T \mathbf{X} \mathbf{X}^T \mathbf{w} = \mathbf{v}^T \mathbf{Y} \mathbf{Y}^T \mathbf{v} = 1$ and can be further reformulated into the following least-squares optimization problem [26]:

$$\begin{aligned} \mathbf{w}, \mathbf{v} &= \arg\min_{\mathbf{w}, \mathbf{v}} \frac{1}{2} \|\mathbf{w}^T \mathbf{X} - \mathbf{v}^T \mathbf{Y}\|_2^2 \\ &\quad s.t. \ \|\mathbf{w}\|_2 = \|\mathbf{v}\|_2 = 1. \end{aligned} \tag{7}$$

To solve the problem in (7), the alternating least squares (ALS) algorithm [33] can be adopted.

According to the aforementioned relationship between CCA and least squares, we reformulate the proposed MCCA into the following least-squares formulation:

$$\begin{aligned} \mathbf{w}_1, \mathbf{w}_3, \mathbf{v} &= \arg\min_{\mathbf{w}_1, \mathbf{w}_3, \mathbf{v}} \frac{1}{2} \|\mathcal{X} \times_1 \mathbf{w}_1^T \times_3 \mathbf{w}_3^T - \mathbf{v}^T \mathbf{Y}\|_2^2 \\ &\quad s.t. \ \|\mathbf{w}_1\|_2 = \|\mathbf{w}_3\|_2 = \|\mathbf{v}\|_2 = 1. \end{aligned} \tag{8}$$

The MCCA in the least-squares framework allows us to impose various regularization operators on it more tractably for different applications. Typically, Tikhonov regularization (L2-regularization) [34], least absolute shrinkage selection operator (LASSO) (L1-regularization) [27] and elastic net (L1- and L2-regularization) [35] can be used, in which the L1-regularization is most applied to learn sparse projection vector for dimensionality reduc-

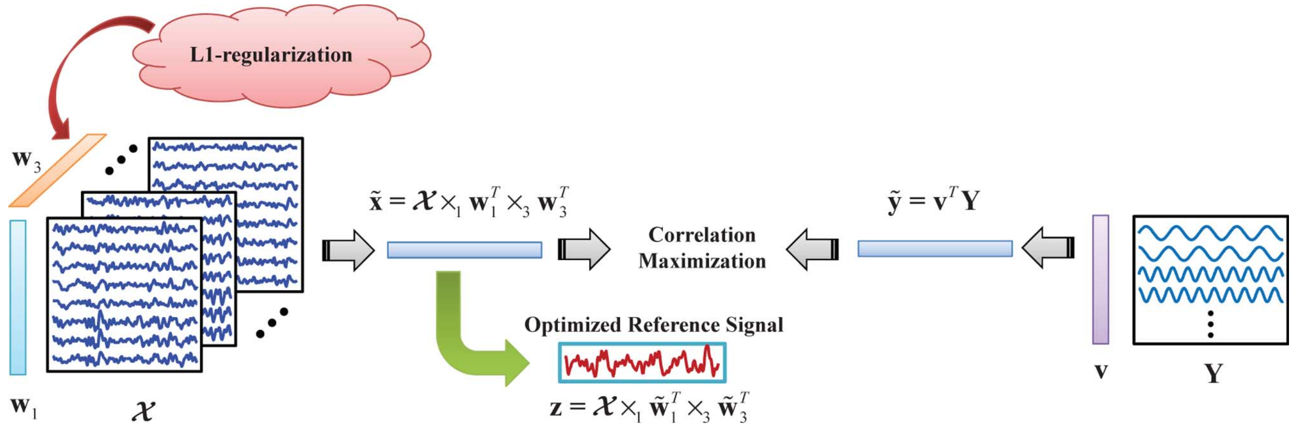


Fig. 2. Illustration of L1-MCCA for reference signal optimization in SSVEP recognition. $\mathcal{X}$ and $\mathbf{Y}$ denote a constructed three-way EEG tensor (channel $\times$ time $\times$ trial) and an original reference signal of sine-cosine (harmonic $\times$ time), respectively. The optimized reference signal $\mathbf{z}$ is obtained by the projection of $\mathcal{X}$ at the channel-way and trial-way with the projection vectors $\tilde{\mathbf{w}}_1$ and $\tilde{\mathbf{w}}_3$ that are learned from the L1-MCCA between $\mathcal{X}$ and $\mathbf{Y}$.

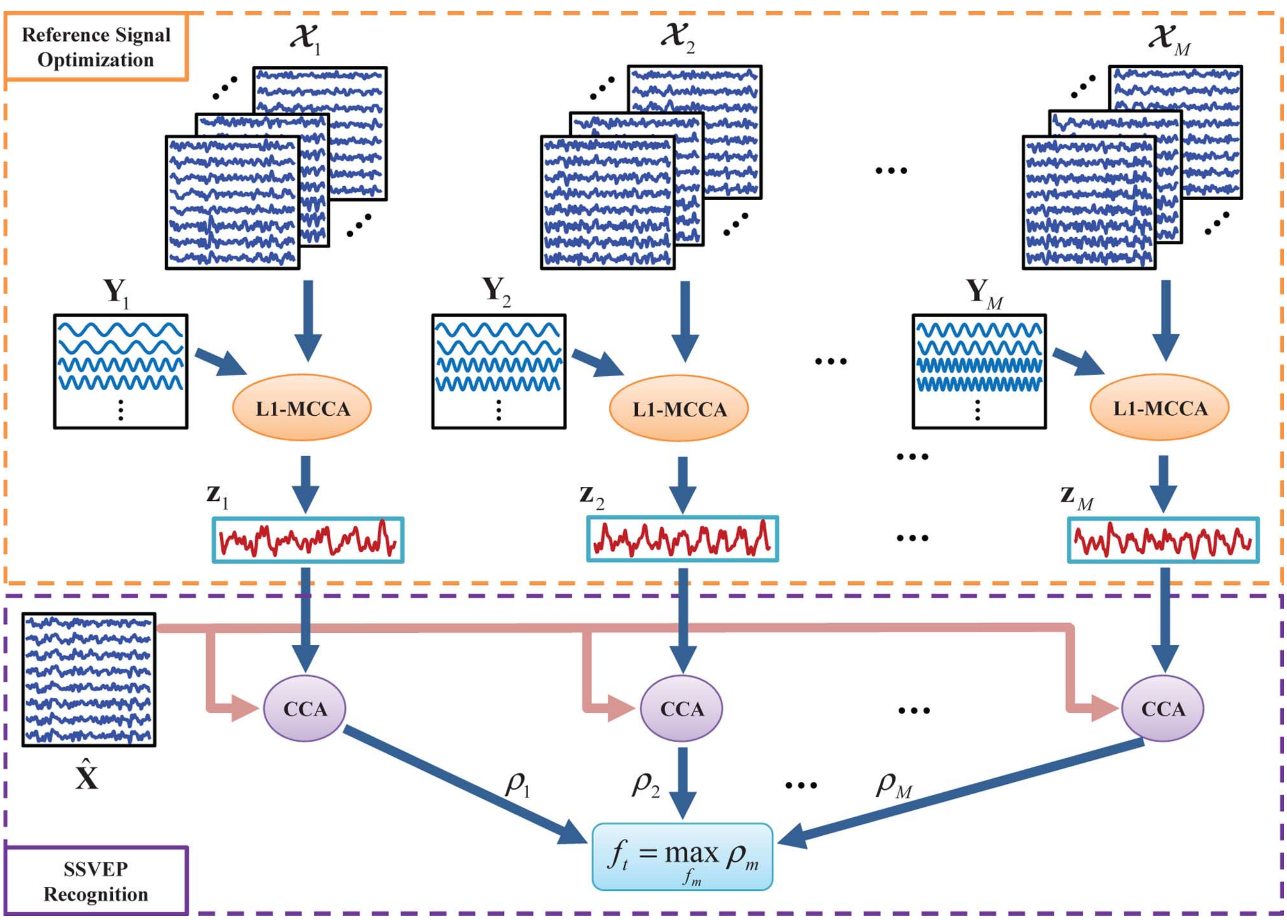


Fig. 3. Illustration of SSVEP recognition model based on the L1-MCCA. Assume three-way tensors (channel $\times$ time $\times$ trial) $\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_M$ are constructed by the EEG data recorded from the experimental runs corresponding to the different M target frequencies, $f_1, f_2, \ldots, f_M$, respectively. After learning the optimal reference signals $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_M$ by L1-MCCAs between the EEG tensors and the sine-cosine signals (harmonic $\times$ time) $\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_M$ with corresponding frequencies, the SSVEP target frequency f_t of a new test data of single trial $\hat{\mathbf{X}}$ is recognized according to the maximum among correlation coefficients $\rho_1, \rho_2, \ldots, \rho_M$ estimated from CCAs between $\hat{\mathbf{X}}$ and the optimized reference signals.

tion and feature selection. With the L1-regularization, a penalized version of the MCCA, called L1-MCCA, is proposed as

$$\begin{aligned} \mathbf{w}_1, \mathbf{w}_3, \mathbf{v} = \arg\min_{\mathbf{w}_1,\mathbf{w}_3,\mathbf{v}} & \frac{1}{2}\|\mathcal{X} \times_1 \mathbf{w}_1^T \times_3 \mathbf{w}_3^T - \mathbf{v}^T\mathbf{Y}\|_2^2 \\ & + \lambda_1\|\mathbf{w}_1\|_1 + \lambda_2\|\mathbf{v}\|_1 + \lambda_3\|\mathbf{w}_3\|_1 \\ s.t.\ \|\mathbf{w}_1\|_2 = \|\mathbf{w}_3\|_2 = \|\mathbf{v}\|_2 = 1 & \end{aligned} \tag{9}$$

where λ_1, λ_2, λ_3 are regularization parameters to control the sparsity of $\mathbf{w}_1$, $\mathbf{v}$, and $\mathbf{w}_3$, respectively. Since the optimization problem in (9) is equivalent to the LASSO problem when any two of $\mathbf{w}_1$, $\mathbf{v}$, and $\mathbf{w}_3$ are fixed, an alternating sparse solution approach is introduced to it, in which $\mathbf{w}_3$ is first fixed to solve $\mathbf{w}_1$ and $\mathbf{v}$ by alternatingly applying the LASSO, then $\mathbf{w}_1$ and $\mathbf{v}$ are fixed to solve $\mathbf{w}_3$ by apply the LASSO again. To solve

Algorithm 2: L1-MCCA algorithm for reference signal optimization in SSVEP recognition

Input: EEG tensor $\mathcal{X} \in \mathbb{R}^{I\times J\times K}$ (channel $\times$ time $\times$ trial) recorded at a specific stimulus frequency, sine-cosine signals $\mathbf{Y} \in \mathbb{R}^{2H\times J}$ (harmonic $\times$ time) constructed as Eq. (2), and regularization parameters λ_1, λ_2 and λ_3.
Output: Optimized reference signal $\mathbf{z} \in \mathbb{R}^J$

Random initialization for $\mathbf{w}_3 \in \mathbb{R}^K$;
Do $\tilde{\mathbf{X}} = \mathcal{X} \times_3 \mathbf{w}_3^T$;
repeat
  Random initialization for $\mathbf{w}_1 \in \mathbb{R}^I$;
  Do $\tilde{\mathbf{x}} = \mathbf{w}_1^T\tilde{\mathbf{X}}$;
  repeat
    $\mathbf{v} = \arg\min_{\mathbf{v}} \frac{1}{2}\|\tilde{\mathbf{x}} - \mathbf{v}^T\mathbf{Y}\|_2^2 + \lambda_2\|\mathbf{v}\|_1$;
    $\mathbf{v} = \mathbf{v}/\|\mathbf{v}\|_2$, $\tilde{\mathbf{y}} = \mathbf{v}^T\mathbf{Y}$;
    $\mathbf{w}_1 = \arg\min_{\mathbf{w}_1} \frac{1}{2}\|\mathbf{w}_1^T\tilde{\mathbf{X}} - \tilde{\mathbf{y}}\|_2^2 + \lambda_1\|\mathbf{w}_1\|_1$;
    $\mathbf{w}_1 = \mathbf{w}_1/\|\mathbf{w}_1\|_2$, $\tilde{\mathbf{x}} = \mathbf{w}_1^T\tilde{\mathbf{X}}$;
  until *Stop criterion is met*;
  $\tilde{\mathbf{X}} = \mathcal{X} \times_1 \mathbf{w}_1^T$;
  $\mathbf{w}_3 = \arg\min_{\mathbf{w}_3} \frac{1}{2}\|\mathbf{w}_3^T\tilde{\mathbf{X}} - \tilde{\mathbf{y}}\|_2^2 + \lambda_3\|\mathbf{w}_3\|_1$;
  $\mathbf{w}_3 = \mathbf{w}_3/\|\mathbf{w}_3\|_2$, $\tilde{\mathbf{X}} = \mathcal{X} \times_3 \mathbf{w}_3^T$;
until *Stop criterion is met*;
$\tilde{\mathbf{w}}_1 = \mathbf{w}_1$, $\tilde{\mathbf{w}}_3 = \mathbf{w}_3$;
$\mathbf{z} = \mathcal{X} \times_1 \tilde{\mathbf{w}}_1^T \times_3 \tilde{\mathbf{w}}_3^T$.

the LASSO problem, this study adopts the coordinate descent algorithm [36] that provides fast and accurate estimation for the sparse solution. The proposed L1-MCCA algorithm for reference signal optimization in SSVEP recognition is summarized in Algorithm 2. After the optimized reference signals are obtained, the SSVEP target frequency of a new test data is recognized by using (1) and (3) again. Figs. 2 and 3 illustrate the reference signal optimization based on the proposed L1-MCCA and the corresponding SSVEP recognition model, respectively.

In the L1-MCCA, L1-regularizations on $\mathbf{w}_1$, $\mathbf{v}$, and $\mathbf{w}_3$ provide automatic selection of channels, harmonics and trials, respectively, for the reference signal optimization. Since the effective channel (eight channels P7, P3, Pz, P4, P8, O1, Oz, and O2 covering the occipital and parietal areas) and harmonic (fundamental frequency component) configurations have been decided by the reliable prior neurophysiologic knowledge [11], [16], only effective trial selection by the L1-regularization is considered in this study, i.e., $\lambda_1 = \lambda_2 = 0$. The sub-loop for $\mathbf{w}_1$ and $\mathbf{v}$ learning in Algorithm 2 can therefore be simply replaced by the ordinary CCA.

III. Experimental Study

A. Experimental Evaluation

The proposed MCCA and L1-MCCA are compared to the ordinary CCA-based method for SSVEP recognition. For both of the MCCA and L1-MCCA, leave-one-run-out cross-validation is implemented to evaluate the average recognition accuracy. More specifically, the data from 19 runs (four trials in each

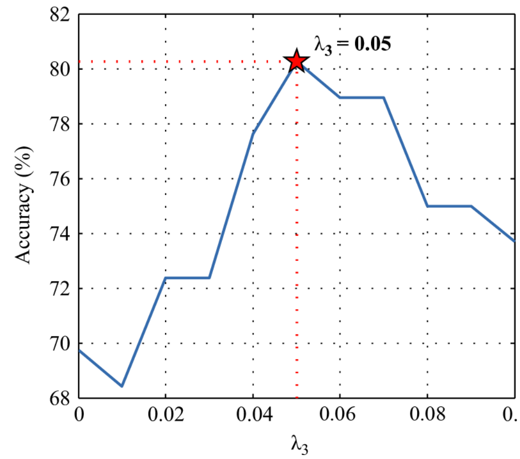


Fig. 4. Example for subject-specific λ_3 selection in the L1-MCCA. The best accuracy is achieved at $\lambda_3 = 0.05$ for S4 with time window length of 1 s. The accuracy at $\lambda_3 = 0$ is obtained by the MCCA.

run) are used as training data for reference signal optimization while the data from the left-out run are used for validation. This procedure is repeated till each run serves once for validation. The optimal subject-specific regularization parameter λ_3 in the L1-MCCA is automatically chosen from a set of values $\lambda_3 \in \{0.01, 0.02, \ldots, 0.1\}$ based on the leave-one-run-out cross-validation during training procedure. As an example, Fig. 4 shows effects of varying λ_3 on the accuracy for S4. The λ_3 value of 0.05 results in the best accuracy on training data, which is used for validation in the subsequent test procedure. Since the CCA method does not require training data for reference signal optimization, the average recognition accuracy is evaluated on the direct validation of 20 runs.

B. Results

Figs. 5 and 6 depict the accuracy for each of the 10 subjects and the average accuracy obtained by the CCA, MCCA, and L1-MCCA, respectively, with different TWs from 1 to 4 s. Table I shows the results of statistical analysis on accuracy differences. These results demonstrate that the proposed MCCA and L1-MCCA significantly outperformed the CCA for SSVEP recognition at all of the four TWs. With L1-regularization on effective trial selection for reference signal optimization, the L1-MCCA performed best and achieved significantly higher accuracy than that of the MCCA at all of the four TWs.

IV. Discussion

Different from the standard sine-cosine waves including no subject-specific and inter-trial information, the reference signals learned from the collaboratively multiway optimization of the MCCA and L1-MCCA contain rich features estimated from the training data, which could actually be referred as a sophisticated calibration procedure. Such features assist to yield more discriminative information for the SSVEP recognition. Fig. 7 shows the correlation coefficients corresponding to different reference signal frequencies (6, 8, 9, and 10 Hz) derived from the leave-one-run-out cross-validation of S10 by the CCA, MCCA, and L1-MCCA, respectively, when each of the four stimulus frequencies was used as the target frequency. Compared to the

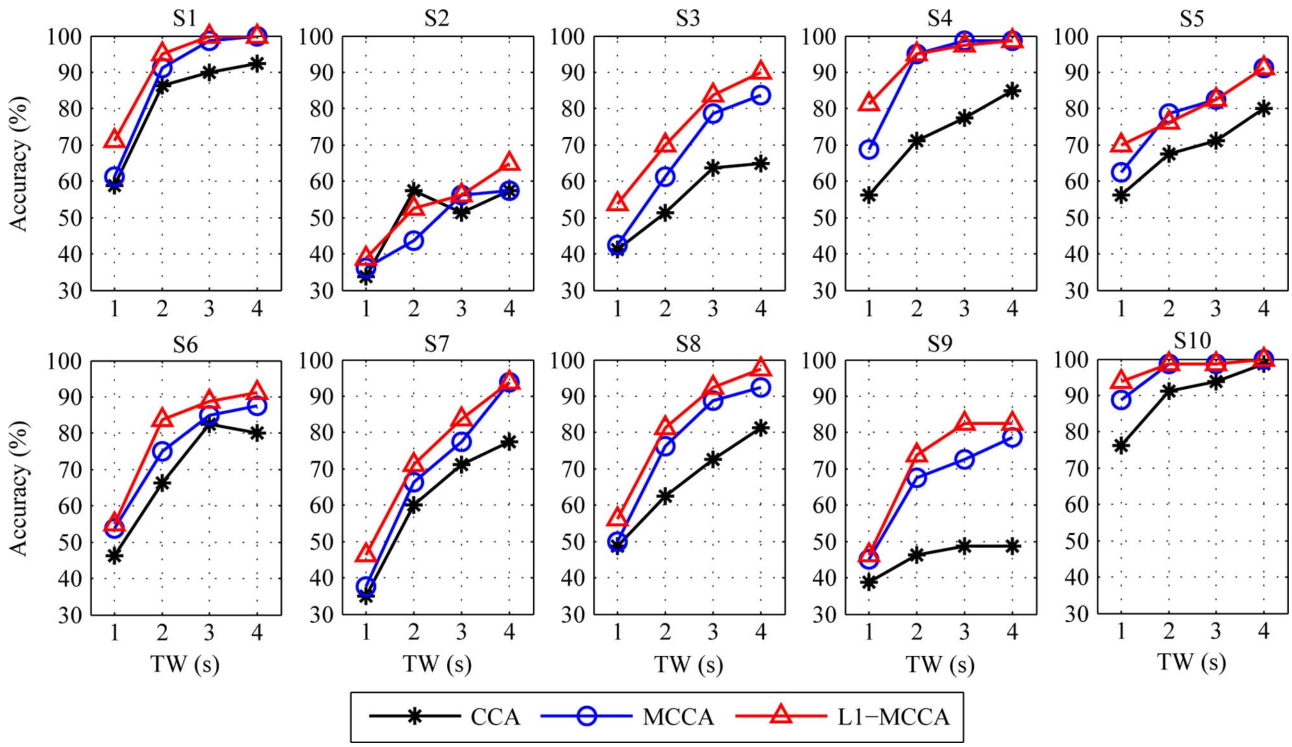


Fig. 5. SSVEP recognition accuracies of the 10 subjects derived by the CCA, MCCA, and L1-MCCA, respectively, with different time window lengths (TW) from 1 to 4 s. Here, the number H of harmonics for sine-cosine waves construction was set to one.

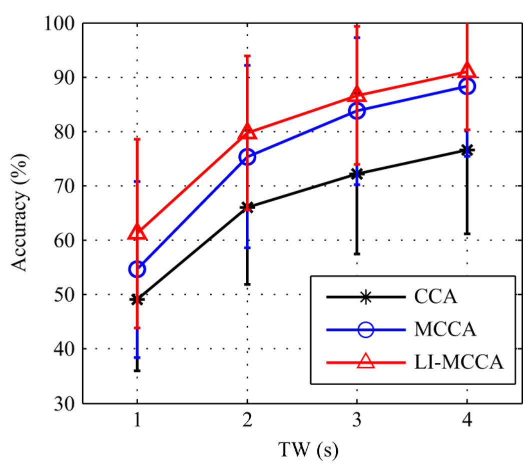


Fig. 6. Averaged SSVEP recognition accuracy derived by the CCA, MCCA, and L1-MCCA, respectively, with different time window lengths (TW) from 1 to 4 s.

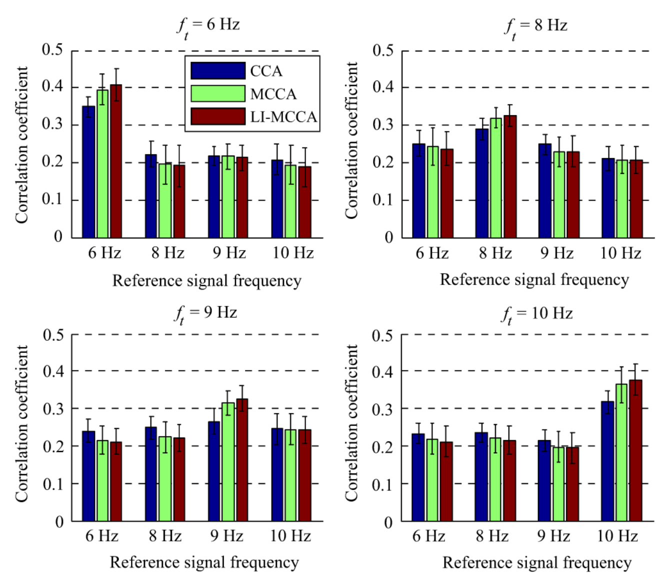


Fig. 7. Correlation coefficients corresponding to different reference signal frequencies (6, 8, 9, and 10 Hz), derived from the leave-one-run-out cross-validation of S10 with time window length of 1 s by the CCA, MCCA, and L1-MCCA, respectively, when each of the four frequencies was used as the target frequency f_t.

TABLE I
STATISTICAL ANALYSIS OF ACCURACY DIFFERENCE BETWEEN EACH TWO METHODS OF THE CCA, MCCA AND, L1-MCCA BY THE PAIRED T-TEST

Method Comparison	Time window length (TW)			
	1 s	2 s	3 s	4 s
MCCA vs. CCA	†	*	††	†
L1-MCCA vs. CCA	§	†	††	††
L1-MCCA vs. MCCA	††	**	*	*

Note: $*p < 0.05$, $**p < 0.01$, $\dagger p < 0.005$, $\dagger\dagger p < 0.001$, $\S p < 0.0001$

CCA, both of the MCCA and L1-MCCA enhanced the correlation coefficient for the target frequency while decreasing those for the nontarget frequencies, which provides an evidence for the superior SSVEP recognition accuracies obtained by the MCCA and L1-MCCA over the CCA. With the L1-regularization on the trial-way array optimization of MCCA for effective trial selection, the obtained L1-MCCA strengthened further the discriminability among correlation coefficients due to the improved reference signals. Fig. 8 depicts the learning results of

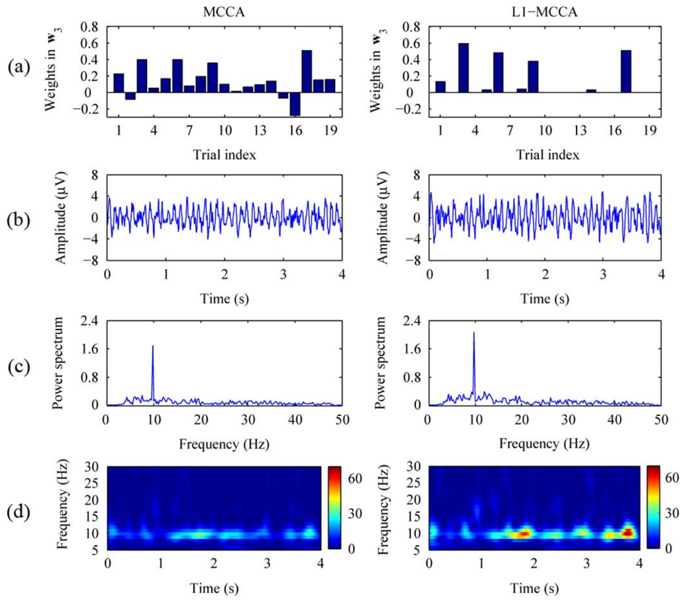


Fig. 8. Learning results of the reference signals at 10 Hz derived from the EEG data of S3 by the MCCA and L1-MCCA, respectively. (a) Weights in $\mathbf{w}_3$, (b) Temporal waveform. (c) Power spectrum obtained by FFT. (d) Time-frequency information obtained by Morlet wavelet transform.

reference signals derived by the MCCA and L1-MCCA, respectively. More sparse trial-way projection vector was obtained by the L1-MCCA compared to the MCCA, which more accurately captured the significant trials including stronger characteristic SSVEP.

The trial-way array optimization in the MCCA and the L1-MCCA is to enhance the SNR by a linear combination of all trials at the corresponding target frequency. Effective implementation of the optimization requires the same or at least similar SSVEP phases of all the trials. In fact, SSVEP has been demonstrated to be phase-locked to the onset of stimulus flicker and there exists a fixed SSVEP phase for a certain stimulus frequency on the same subject [17], [37], [38]. Since the EEG segments extracted in our study always initiate at the onset of stimulus flicker, the requirement of phase consistency was therefore met for the MCCA and the L1-MCCA. To confirm this point, we show an example of the SSVEP phases at different target frequencies (see Fig. 9). The phase at target frequency f_t is computed by

$$\phi(f_t) = \arctan\left(\frac{\mathrm{im}(f_t)}{\mathrm{re}(f_t)}\right), -360^\circ < \phi(f_t) \leqslant 0 \quad (10)$$

where $\mathrm{im}(f_t)$ and $\mathrm{re}(f_t)$ denote the imaginary and real parts of FFT, respectively, on the single-channel EEG data at f_t. The negativity in phase indicates that the SSVEP lags behind the onset of stimulus flicker. For each target frequency, the SSVEP phases of all trials are concentrated and present the phase-locking property. Through the trial-way array optimization exploiting the phase-locking property of SSVEP, the MCCA and the L1-MCCA learned more effective reference signals from multiple trials to improve the recognition performance.

Since the MCCA and the L1-MCCA algorithms involve an alternating iteration procedure, an appropriate stop criterion should be predefined to guarantee convergence of the iteration

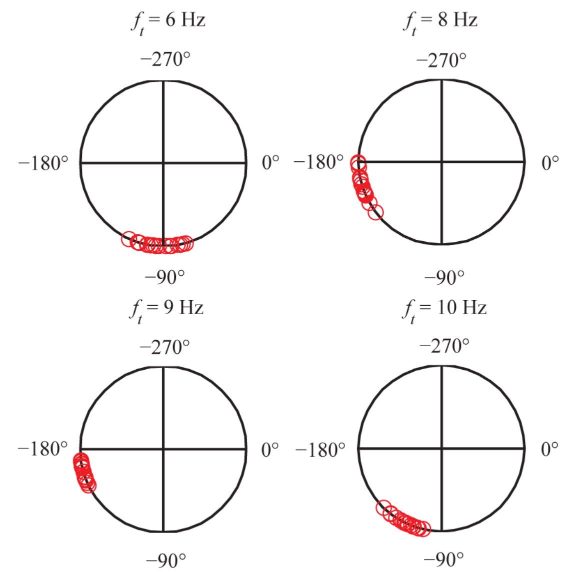


Fig. 9. SSVEP phases at different target frequencies estimated by FFT from the 20 runs (i.e., 20 trials for each target frequency) of S1. Here, channel Oz was used for phase estimation. The phases are plotted in red marks on unit circles.

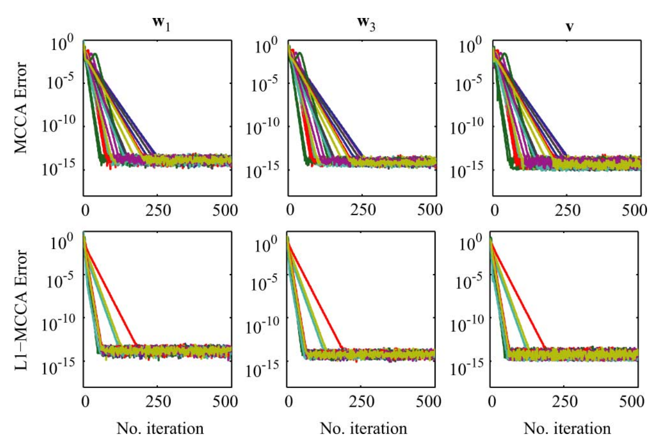


Fig. 10. Convergence curves of the projection vectors $\mathbf{w}_1$, $\mathbf{w}_3$, and $\mathbf{v}$ obtained during the leave-one-run-out cross-validation of S3 by the MCCA and L1-MCCA, respectively (Oscillations in the convergence curves may be caused by the machine precision in Matlab).

process. As an example, 20 convergence curves derived during the leave-one-run-out cross-validation of S3 are plotted for each of $\mathbf{w}_1$, $\mathbf{w}_3$, and $\mathbf{v}$ in the MCCA and L1-MCCA, respectively (see Fig. 10). The results show that stable convergences (iteration error $< 10^{-13}$) of the three projection vectors are achieved for the MCCA within 260 iteration steps and achieved for the L1-MCCA within 200 iteration steps. To achieve a trade-off between efficiency and accuracy, the stop criterion used in this study is defined as

$$\mathrm{Error} = \|\mathbf{w}(n) - \mathbf{w}(n-1)\|_2 < 10^{-5} \quad (11)$$

where n denotes the number of iteration steps, $\mathbf{w}$ is the projection vector to be learned. That is, The iteration process does not stop until each of $\mathbf{w}_1$, $\mathbf{w}_3$, and $\mathbf{v}$ satisfies (11).

To comprehensively assess efficiencies of the proposed methods, Table II investigates the computational time of the MCCA and the L1-MCCA for reference signal optimization,

TABLE II
COMPUTATIONAL TIME OF THE MCCA AND THE L1-MCCA FOR REFERENCE SIGNAL OPTIMIZATION, AND THAT OF THE CCA FOR SSVEP RECOGNITION, EVALUATED FROM THE LEAVE-ONE-RUN-OUT CROSS-VALIDATION OF S3

Method	Time window length (TW)			
	1 s	2 s	3 s	4 s
MCCA	1.22 s	1.46 s	1.49 s	1.75 s
L1-MCCA	1.58 s	2.05 s	2.37 s	2.39 s
CCA	0.0045 s	0.0046 s	0.0052 s	0.0056 s

and that of the CCA for SSVEP recognition. The computation environment is under Matlab R2011a on a laptop with 1.20 GHz CPU (3 GB RAM). The reference signal optimization procedures of the MCCA and the L1-MCCA can be executed in 1.75 and 2.39 s, respectively, using a TW of 4 s. The computational cost is negligible in contrast to the time cost for training data recording. After reference signal optimization, the SSVEP recognition procedures of the MCCA and the L1-MCCA based on the CCA can be executed in 0.0056 s using a TW of 4 s. Thus, computational efficiencies of the MCCA and the L1-MCCA are sufficiently high for practical application.

Collaboratively multiway optimization has been suggested to be more promising than the one-way optimization for EEG data analysis [39]–[43] and also for electrocorticogram (ECoG) data analysis [44], [45]. Recently, regularized tensor discriminant analysis and penalized N-way partial least squares have began to emerge and shown great potentials for the BCI application [46], [47]. In this study, the proposed L1-MCCA effectively combines both the multiway canonical correlation analysis and L1-regularization technique, which significantly improves the classification accuracy of SSVEP-based BCI. Thus, it can be seen that an ingenious combination of multiway analysis and regularization could be a promising approach for the BCI application, which is worthy of our further studies.

Note that the accuracies: 49.2% with a 1 s time window length (TW), 66.0% with a 2 s TW, 72.3% with a 3 s TW, and 76.6% with a 4 s TW averaged on 10 subjects for 4-class SSVEP classification obtained by the CCA in our study, are lower than the results reported in some general literatures. For example, Guger *et al.* reported 95.5% accuracy with a 3 s TW averaged on 53 subjects for 4-class SSVEP classification [48]. See also the results reported in [11] and [16]. However, it has been reported that some subjects obtained significantly poor accuracies due to their annoyances and fatigues on flickering stimuli, and hence failed to use the SSVEP-based BCI [8], [49]. In our study, by using the CCA, accuracy of the worst subject was really low (lower than 40% with a 1 s TW and lower than 50% with a 4 s TW) whereas that of the best subject was relatively high (higher than 76% with a 1 s TW and higher than 98% with a 4 s TW). Therefore, the average accuracy depends on the subject group. Actually, several literatures also reported similar average results as our study. Müller–Putz *et al.* reported 63.9% accuracy with a 2 s TW averaged on nine subjects for 4-class SSVEP classification [9] and reported 72.5% accuracy with a 4 s TW averaged on four subjects for 4-class SSVEP classification [50]. Pan *et al.* reported 46.2% accuracy with a 1 s TW, 59.2% accuracy with a 2 s TW, 68.8% accuracy with a 3 s TW, and 73.8% accuracy with a 4 s TW averaged on 10 subjects for 6-class SSVEP classification by using the CCA [17].

V. CONCLUSION

This study introduced an L1-regularized multiway canonical correlation analysis (L1-MCCA) to enhance SSVEP recognition accuracy for the BCI application. Collaborative optimization based on the CCA was implemented on the channel-way and trial-way arrays of the constructed multi-channel and multi-trial EEG tensor to learn the reference signals in correlation analysis for SSVEP recognition. L1-regularization was further imposed on the trial-way array optimization for effective trial selection to improve the reference signals. Experimental results of 10 subjects demonstrated both the proposed MCCA and L1-MCCA achieved significantly higher SSVEP recognition accuracies than that of the standard CCA without reference signal optimization. Furthermore, the accuracy is further improved by the L1-MCCA in comparison with the MCCA. Future studies will exploit other types of regularization on the MCCA and investigate their effectiveness on performance improvement of the SSVEP-based BCI.

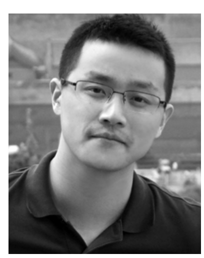
ACKNOWLEDGMENT

The authors would like to thank the Associate Editor and the anonymous reviewers for their insightful comments and suggestions that helped improve the paper.

REFERENCES

[1] J. Wolpaw, N. Birbaumer, D. McFarland, G. Pfutscheller, and T. Vaughan, "Brain-computer interfaces for communication and control," *Clin. Neurophysiol.*, vol. 113, no. 6, pp. 767–791, 2002.
[2] Y. Zhang, Q. Zhao, J. Jin, X. Wang, and A. Cichocki, "A novel BCI based on ERP components sensitive to configural processing of human faces," *J. Neural Eng.*, vol. 9, no. 2, pp. 026018–026018, 2012.
[3] Y. Li, J. Long, T. Yu, Z. Yu, C. Wang, H. Zhang, and C. Guan, "An EEG-based BCI system for 2-D cursor control by combining mu/beta rhythm and P300 potential," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 10, pp. 2495–2505, Oct. 2010.
[4] G. Pfurtscheller, C. Brunner, A. Schlögl, and S. Lopes, "Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks," *NeuroImage*, vol. 31, no. 1, pp. 153–159, 2006.
[5] X. Gao, D. Xu, M. Cheng, and S. Gao, "A BCI-based environmental controller for the motion-disabled," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, no. 2, pp. 137–140, Jun. 2003.
[6] G. Pfurtscheller, T. Solis-Escalante, R. Ortner, P. Linortner, and G. Müller-Putz, "Self-paced operation of an SSVEP-based orthosis with and without an imagery-based brain switch: A feasibility study towards a hybrid BCI," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 18, no. 4, pp. 409–414, Aug. 2010.
[7] 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.
[8] B. Allison, D. McFarland, G. Schalk, S. Zheng, M. Jackson, and J. Wolpaw, "Towards an independent brain-computer interface using steady state visual evoked potentials," *Clin. Neurophysiol.*, vol. 119, no. 2, pp. 399–408, 2008.
[9] G. Müller-Putz, R. Scherer, C. Brauneis, and G. Pfurtscheller, "Steady-state visual evoked potential (SSVEP)-based communication: Impact of harmonic frequency components," *J. Neural Eng.*, vol. 2, no. 4, pp. 123–130, 2005.
[10] 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, Oct. 2002.
[11] 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, Dec. 2006.
[12] 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, Apr. 2007.

[13] C. Wu, H. Chang, P. Lee, K. Li, J. Sie, C. Sun, C. Yang, P. Li, H. Deng, and K. Shyu, “Frequency recognition in an SSVEP-based brain computer interface using empirical mode decomposition and refined generalized zero-crossing,” *J. Neurosci. Methods*, vol. 196, no. 1, pp. 170–181, 2011.
[14] Y. Zhang, G. Zhou, Q. Zhao, A. Onishi, J. Jin, X. Wang, and A. Cichocki, “Multiway canonical correlation analysis for frequency components recognition in SSVEP-based BCIS,” in *Proc. 18th Int. Conf. Neural Inf. Process.*, 2011, pp. 287–295.
[15] Z. Wu and D. Yao, “Frequency detection with stability coefficient for steady-state visual evoked potential (SSVEP)-based BCIS,” *J. Neural Eng.*, vol. 5, no. 1, pp. 36–43, 2008.
[16] G. Bin, X. Gao, Z. Yan, B. Hong, and S. Gao, “An online multi-channel SSVEP-based brain-computer interface using a canonical correlation anaylsis method,” *J. Neural Eng.*, vol. 6, no. 4, p. 046002, 2009.
[17] 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, pp. 036027–036027, 2011.
[18] 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, p. e29519, 2012.
[19] J. Xie, G. Xu, J. Wang, F. Zhang, and Y. Zhang, “Steady-state motion visual evoked potentials produced by oscillating newton’s rings: Implications for brain-computer interfaces,” *PLoS One*, vol. 7, no. 6, p. e39707, 2012.
[20] E. Yin, Z. Zhou, J. Jiang, F. Chen, Y. Liu, and D. Hu, “A novel hybrid BCI speller based on the incorporation of SSVEP into the P300 paradigm,” *J. Neural Eng.*, vol. 10, no. 2, p. 026012, 2013.
[21] 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.
[22] A. Cichocki, R. Zdunek, A.-H. Phan, and S. Amari, in *Nonnegative Matrix and Tensor Factorization: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*, New York, Wiley, 2009.
[23] A. Cichocki, Y. Washizawa, T. Rutkowski, H. Bakardjian, A.-H. Phan, S. Choi, H. Lee, Q. Zhao, L. Zhang, and Y. Li, “Noninvasive BCIs: Multiway signal-processing array decompositions,” *IEEE Computer*, vol. 41, no. 10, pp. 34–42, Oct. 2008.
[24] T. Kim and R. Cipolla, “Canonical correlation analysis of video volume tensors for action categorization and detection,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 8, pp. 1415–1428, Aug. 2009.
[25] D. Tao, X. Li, X. Wu, and S. Maybank, “General tensor discriminant analysis and Gabor features for gait recognition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 10, pp. 1700–1715, Oct. 2007.
[26] D. Hardoon and J. Shawe-Taylor, “Sparse canonical correlation analysis,” *Mach. Learn.*, vol. 83, no. 3, pp. 331–353, 2011.
[27] R. Tibshirani, “Regression shrinkage and selection via the LASSO,” *J. R. Stat. Soc. B*, vol. 58, no. 1, pp. 267–288, 1996.
[28] Y. Li, A. Cichocki, and S. Amari, “Analysis of sparse representation and blind source separation,” *Neural Comput.*, vol. 16, no. 6, pp. 1193–1234, 2004.
[29] H. Hotelling, “Relations between two sets of variates,” *Biometrika*, vol. 28, no. 3/4, pp. 321–377, 1936.
[30] O. Friman, J. Cedefamn, P. Lundberg, M. Borga, and H. Knutsson, “Detection of neural activity in functional MRI using canonical correlation analysis,” *Magn. Reson. Med.*, vol. 45, no. 2, pp. 323–330, 2001.
[31] T. Kolda and B. Bader, “Tensor decompositions and applications,” *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009.
[32] D. Witten and R. Tibshirani, “Penalized classification using Fisher’s linear discriminant,” *J. R. Stat. Soc. B*, vol. 73, no. 5, pp. 753–772, 2011.
[33] F. Yates, “The analysis of replicated experiments when the field results are incomplete,” *Empire J. Exp. Agric.*, vol. 1, no. 2, pp. 129–142, 1933.
[34] A. Tikhonov and V. Arsenin, in *Solutions of Ill-Posed Problems.*, New York, Wiley, 1977.
[35] H. Zou and T. Hastie, “Regularization and variable selection via the elastic net,” *J. R. Stat. Soc. B*, vol. 67, no. 2, pp. 301–320, 2005.
[36] J. Friedman, T. Hastie, and R. Tibshirani, “Regularization paths for generalized linear models via coordinate descent,” *J. Stat. Softw.*, vol. 33, no. 1, pp. 1–22, 2010.
[37] E. Sutter, “The brain response interface: Communication through visually-induced electrical brain responses,” *J. Microcomput. Appl.*, vol. 15, no. 1, pp. 15–31, 1992.
[38] P. Lee, J. Sie, Y. Liu, C. Wu, M. Lee, C. Shu, P. Li, C. Sun, and K. Shyu, “An SSVEP-actuated brain computer interface using phase-tagged flickering sequences: A cursor system,” *Ann. Biomed. Eng.*, vol. 38, no. 7, pp. 2383–2397, 2010.
[39] Y. Zhang, G. Zhou, Q. Zhao, J. Jin, X. Wang, and A. Cichocki, “Spatial-temporal discriminant analysis for ERP-based brain-computer interface,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 21, no. 2, pp. 233–234, Mar. 2013.
[40] J. Li, L. Zhang, D. Tao, H. Sun, and Q. Zhao, “A prior neurophysiologic knowledge free tensor-based scheme for single trial EEG classification,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 17, no. 2, pp. 107–115, Apr. 2009.
[41] W. Wu, Z. Chen, S. Gao, and E. Brown, “A hierarchical Bayesian approach for learning sparse spatio-temporal decompositions of multichannel EEG,” *NeuroImage*, vol. 56, no. 4, pp. 1929–1945, 2011.
[42] F. Cong, A.-H. Phan, P. Astikainen, Q. Zhao, Q. Wu, J. Hietanen, T. Ristaniemi, and A. Cichocki, “Multi-domain feature extraction for small event-related potentials through nonnegative multi-way array decomposition from low dense array EEG,” *Int. J. Neural Syst.*, vol. 23, no. 2, p. 1350006, 2013.
[43] C. Park, D. Looney, R. Naveed, A. Ahrabian, and D. Mandic, “Classification of motor imagery BCI using multivariate empirical mode decomposition,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 21, no. 1, pp. 10–22, Jan. 2013.
[44] A. Eliseyev, C. Moro, T. Costecalde, N. Torres, S. Gharbi, C. Mestais, A. Benabid, and T. Aksenova, “Iterative N-way PLS for self-paced BCI in freely moving animals,” *J. Neural Eng.*, vol. 8, no. 4, pp. 046012–046012, 2011.
[45] Q. Zhao, C. Caiafa, D. Mandic, Z. Chao, Y. Nagasaka, N. Fujii, L. Zhang, and A. Cichocki, “Higher-order partial least squares (HOPLS): a generalized multi-linear regression method,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 7, pp. 1660–1673, Jul. 2013.
[46] A. Eliseyev, C. Moro, J. Faber, N. Torres, C. Mestais, A. Benabid, and T. Aksenova, “L1-penalized N-way PLS for subset of electrodes selection in BCI experiments,” *J. Neural Eng.*, vol. 9, no. 4, pp. 045010–045010, 2012.
[47] J. Li and L. Zhang, “Regularized tensor discriminant analysis for single trial EEG classification in BCI,” *Pattern Recogn. Lett.*, vol. 31, no. 7, pp. 619–628, 2010.
[48] C. Guger, B. Allison, B. Großwindhager, R. Prückl, C. Hintermüller, C. Kapeller, M. Bruckner, G. Krausz, and G. Edlinger, “How many people could use an SSVEP BCI?,” *Front. Neurosci.*, vol. 6, p. 169, 2012.
[49] B. Allison, T. Luth, D. Valbuena, A. Teymourian, I. Volosyak, and A. Graser, “Demographics: How many (and what kinds of) people can use an SSVEP BCI?,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 18, no. 2, pp. 107–116, Apr. 2010.
[50] G. Müller-Putz and G. Pfurtscheller, “Control of an electrical prosthesis with an SSVEP-based BCI,” *IEEE Trans. Biomed. Eng.*, vol. 55, no. 1, pp. 361–364, Jan. 2008.



Yu Zhang (M’13) received the Ph.D. degree in control science and engineering from the School of Information Science and Engineering, East China University of Science and Technology, Shanghai, China, in 2013.

He worked as an International Program Associate (IPA) for two years (from December 2010 to November 2012) in the Laboratory for Advanced Brain Signal Processing (LABSP) at RIKEN Brain Science Institute, Wako-shi, Japan. He is currently an Assistant Professor at the School of Information Science and Engineering, East China University of Science and Technology, Shanghai, China. His research interests include brain–computer interface, signal processing, tensor analysis, machine learning, and pattern recognition.

Guoxu Zhou (M’12) was born in Hubei Province, China, in 1977. He received the Ph.D. degree in intelligent signal and information processing from South China University of Technology, Guangzhou, China, in 2010.

He is currently a Research Scientist of the laboratory for Advanced Brain Signal Processing, RIKEN Brain Science Institute, Wako-shi, Japan. His research interests include statistical signal processing, tensor analysis, intelligent information processing, and machine learning.

Jing Jin received the Ph.D. degree in control theory and control engineering from the East China University of Science and Technology, Shanghai, China, in 2010. His Ph.D. advisors were Prof. Gert Pfurtscheller at Graz University of Technology from 2008 to 2010 and Prof. Xingyu Wang at East China University of Science and Technology from 2006 to 2008.

He is currently an Associate Professor at East China University of Science and Technology. His research interests include brain–computer interface, signal processing, and pattern recognition.

Minjue Wang received the B.S. degree in control technology and instrument, in 2011, from the School of Information Science and Engineering at East China University of Science and Technology, Shanghai, China, where he is currently working toward the M.S. degree.

His current research interests include brain–computer interface, machine learning, and pattern recognition.

Xingyu Wang was born in Sichuan, China, in 1944. He received the B.S. degree in mathematics from Fudan University, Shanghai, China, in 1967, the M.S. degree in control theory from East China Normal University, Shanghai, China, in 1982, and the Ph.D. degree in industrial automation from East China University of Science and Technology, Shanghai, China, in 1984.

He is currently a Professor at the School of Information Science and Engineering, East China University of Science and Technology, Shanghai, China. His research interests include control theory, control techniques, the application to biomedical system, and brain control.

Andrzej Cichocki (F'13) received the M.Sc. (with Hons.), Ph.D., and Dr.Sc. (Habilitation) degrees, all in electrical engineering, from Warsaw University of Technology, Warsaw, Poland.

Since 1972, he has been with the Institute of Theory of Electrical Engineering, Measurement and Information Systems, Faculty of Electrical Engineering, Warsaw University of Technology, Warsaw, Poland, where he received the title of a Full Professor in 1995. He spent several years at the University Erlangen–Nuerenberg, Germany, at the Chair of Applied and Theoretical Electrical Engineering directed by Prof. R. Unbehauen, as an Alexander-von-Humboldt Research Fellow and Guest Professor. From 1995 to 1997, he was a team leader of the laboratory for Artificial Brain Systems, at the Frontier Research Program RIKEN, Japan, in the Brain Information Processing Group. He is currently the Head of the laboratory for Advanced Brain Signal Processing, at RIKEN Brain Science Institute, Wako-shi, Japan. He is author of more than 250 technical papers and four monographs (two of which have been translated to Chinese). His research interests include signal processing, inverse problems, neural network and learning algorithms, tensor analysis, and brain–computer interface.