Optimizing Filter-bank Canonical Correlation Analysis for fast response SSVEP Brain-Computer Interface (BCI)

Aung Aung Phyo Wai[1], Heng Guo[2], Ying Chi[2], Lei Zhang[2], Xian-Sheng Hua[2] and Cuntai Guan[1]
[1]*School of Computer Science and Engineering, Nanyang Technological University, Singapore*
[2]*HealthTech Division, DAMO Academy, Alibaba Group Holding Ltd*
{apwaung, ctguan}@ntu.edu.sg
{gh205191, xinyi.cy, lei.zhang.lz, xiansheng.hxs}@alibaba-inc.com

***Abstract*—Steady-State Visual Evoked Potential (SSVEP) BCI brings high accuracy and consistent performance across subjects at the expense of a long stimulus presentation time window. Several recent methods exploited subject-specific features to improve SSVEP recognition performance in a short time window less than 1s. Although the calibration process is tedious and causes inconvenience, small calibration data with short duration resulting in higher performance gains are worth considering. So we propose a method by optimizing Filter-Bank Canonical Correlation Analysis (FBCCA) with subjects' calibrated templates, subject-specific weights and multiple reference types. The proposed method, subject-calibration extended FBCCA (SCEF) leverages independent and distinct discrimination characteristics of multiple references with subject-specific weight-adjusted features to improve SSVEP recognition performance. We tested the proposed method with different parameters compared with FBCCA baseline and state-of-the-art calibration methods on forty targets SSVEP dataset using 0.2s to 4s time windows. Our evaluation results show SCEF with three reference templates and subject-specific weighted features perform significantly better than all FBCCA variants in 0.2 s to 1 s time window ($p < 0.001$). SCEF performs marginally, not statistically significant, better than existing methods about $2.69 \pm 2.32\%$ mean accuracy across time windows. Including multiple templates and subject-specific weight increases $15.73 \pm 5.34\%$ and $8.06 \pm 2.06\%$ in mean accuracy resulting the overall performance improvements in short time window. The proposed optimization only requires prior calibration data to create subject-specific templates and weights instead of learning features from calibration data every time. This enables not requiring to repeat the calibration step in every SSVEP session for the same subject while still maintaining accuracy similar to state-of-the-art calibration methods.**

***Index Terms*—Brain Computer Interface, Steady-State Visual Evoked Potential, Filter-Bank, Canonical Correlation Analysis, Subject Calibration.**

I. Introduction

Brain-Computer Interface (BCI) enables users to interact with both physical and virtual world directly through brain signals in different means [1]. Depending on the stimuli used, brain's response types and user's participation levels, three broad categories of BCI solutions such as active, reactive and passive BCI. Steady-State Visual Evoked Potentials (SSVEP) is visual reactive BCI modality that exhibits brain responses called Steady-State Responses in response to the stimulus frequencies presented within user's field of view. Steady-State Responses are mixtures of evoked brain signals in response fundamental and harmonic components of stimulus frequency and, ongoing spontaneous brain activities [2]. SSVEP BCI solutions can be ranged from control of robots [3] and communication such as visual speller [4] to vision research [5] and clinical applications [6]. Because of the consistent and objective responses to stimulus presentation, SSVEP might be used as objective assessment method instead of subjective psycho-physic methods to recognize non-healthy eyes in Glaucoma patients [6]. Depending on application specific requirements, we can see different SSVEP applications with diverse characteristics such as stimulus design, stimulus coding, stimulus frequency ranges, presentation paradigm such as flickering or motion reversal, number of targets from one to multiples. These diversified application needs fuel SSVEP research in many areas such as understanding Steady-State Response characteristics [7], improving stimulus design and presentation [8], improving decoding methods [9], [10]. Our work focuses on improving SSVEP decoding in quick response time by exploiting intra- and inter- subject information to create subjects' templates for target detection and adjust subject-specific weights for feature fusion.

Frequency recognition in SSVEP BCI requires either long segment length or subject-specific calibration to achieve high performance [11]. But high throughput and near real-time response performance requires detecting the correct stimulus frequency in quick response time [12]. By exploiting existence of multiple channels localized at occipital areas and robust oscillatory steady-state responses to SSVEP stimuli, SSVEP target recognition achieves high accuracy in a time window longer than one second. Multivariate statistical analysis based Canonical Correlation Analysis (CCA) is commonly used as a baseline method in SSVEP analysis as no calibration or training is required with sufficient high accuracy in a long time window [13]. CCA exhibits high accuracy comparable with other training and calibration methods such as power spectral

Cuntai Guan is the corresponding author. Email: ctguan@ntu.edu.sg

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

density analysis in long data length although having poor accuracy in short data length [11]. Although general subject-independent target classification is desirable, many new algorithms leverage subject-specific calibration to achieve high accuracy in short time window. Various data-driven multivariate methods are proposed to extract unique SSVEP representations from subject-specific calibration data [14], [15]. These methods show tremendous performance improvement in a short time window compared with various extended CCA methods. Furthermore, filter-bank extension of these methods show additional performance gains in their evaluation [10]. Among extended CCA methods, individual-templates based CCA outperforms other CCA methods in various studies by identifying correlation vectors among test data, subject-specific reference and ideal sine-cosine reference [16], [17]. Recently, deep learning approach shows better performance than baseline methods but evaluation only used non-standard SSVEP dataset with time window of 1s only [18]. Cross-subject transfer learning approach aims to reduce calibration efforts that are troublesome or infeasible in practical applications [19]. Still such works are in early stages with further validation required to detect salient SSVEP features in different cross-subjects, cross-days and cross-modalities scenarios [20]. Adaptation of spectrum and phase characteristics of individuals shows higher performance than baseline methods [9]. Either subject-specific adaptation or subject-to-subject transfer learning applied to baseline methods can further improve SSVEP decoding performance. So, our study focuses on assessing the advantages of using multiple reference templates with subject-specified weights adjusted features of filter-bank outputs for improving performance in different time windows ranging from 0.2 to 4s.

So we propose an extension to improve the poor performance of FBCCA in short time window by incorporating subjects' calibration references and subject-specific weights to fuse sub-bands features [21]. As FBCCA enables additional performance gains in different SSVEP recognition methods, our extension of FBCCA enables yet another performance gains by applying to other multivariate methods beside CCA [22]. We also tested the proposed SCEF method with open-source standard 40-class SSVEP data-set using different baseline and state-of-the-art methods. The below section II outlines the detailed description of the proposed method. Section III explains the evaluation data-set with parameters and criteria used in performance analysis. Section IV presents the evaluation results together with discussion and future works followed by conclusion in section V.

II. Proposed Method

CCA uses two multivariate variables such as multi-channel EEG and reference sine-cosine stimulus frequency templates in SSVEP classification to find the maximum correlation from linear transformation with canonical variates [13]. In standard CCA operation, the maximum correlation coefficient ρ between canonical variates $U = X^T W_x$ where $X \in \mathbb{R}^{N_c, N_t, N_d, N_f}$ and $V = Y^T W_y$ where $Y \in \mathbb{R}^{2N_h, N_d, N_c}$ can be derived by maximizing spatial weights W_x and W_y among withing class matrices S_x, S_y and between-class matrix, S_{xy} as below Equation. 1. Here denote that N_c is number of channels, N_t is number of trials, N_d is number of data samples, N_f is number of frequencies or targets and N_h is number of harmonics.

$$\rho = \max_{W_x, W_y} \frac{W_x^T S_{xy} W_y}{\sqrt{W_x^T S_x W_x}\sqrt{W_y^T S_y W_y}} \tag{1}$$

The target frequency can be identified by selecting the frequency with the maximum correlation among all stimuli as shown in Equation. 2.

$$f_{\text{CCA}} = \max_f \rho(f), \ f = f_1, f_2, \ldots, f_{N_f} \tag{2}$$

Because of resulting poor SSVEP detection accuracy in short time window, several training-free and training-based extensions of CCA are proposed to further improve the frequency recognition accuracy [16], [17], [21]. Among these extensions, Filter-Bank CCA (FB-CCA) requires no subject specific calibration data though optimal sub-band weight coefficients must be pre-defined for each data-set and method [21], [22]. In this method, multi-channels EEG signals X are decomposed into multiple N sub-bands through pre-defined zero-phase band-pass filters resulting X_{SB}^n components [21]. Then, CCA is applied to each sub-band with the ideal Sine-Cosine (SC) reference signals Y_{SC} to compute correlation vector ρ_f^n. The target frequency, f_{FBCCA} can be identified as the maximum of $\hat{\rho_f}$ in Equation. 2 that are weighted sum of the square of the correlation coefficients from all N sub-band components.

$$\hat{\rho_f} = \sum_{n=1}^{N} w_n.(\rho_f^n)^2. \tag{3}$$

The weight w_n is defined as $w_n = n^{(-a)} + b$ by weighting scores from n sub-bands according to the known Steady-State Response characteristics of decreasing in response strength of harmonic components with respect to increasing target stimulus frequency. We used n sub-bands where each sub-band starts with unique m frequencies but ends with about approximately five or six times of f_{max} of stimulus frequencies. In original FBCCA analysis, a and b constants of weight vector w_n are set as 1.25 and 0.25 respectively after parameters grid search [21]. The number of FB (n) is selected at 5 with the starting frequency for n^{th} sub-bands is set at $n \times 8$ Hz with same ending frequency at 88Hz [14], [21].

Filter-bank as pre-processing step in different training-free and training-based baseline methods shows substantial performance improvements in SSVEP frequency recognition [22]. The proposed method, Subject Calibration Extended FB-CCA (SCEF), as shown in Figure. 1 requires subjects' calibration data to create subject-specific and subject-independent reference templates [17]. In offline phase, calibration EEG data, X_{train}, are decomposed into n sub-bands of X_{test}^{SB} using pre-defined band-pass filters [21] to create both Subject-Specific (SS) templates, Y_{SS}, and Subject-Independent (SI) templates,

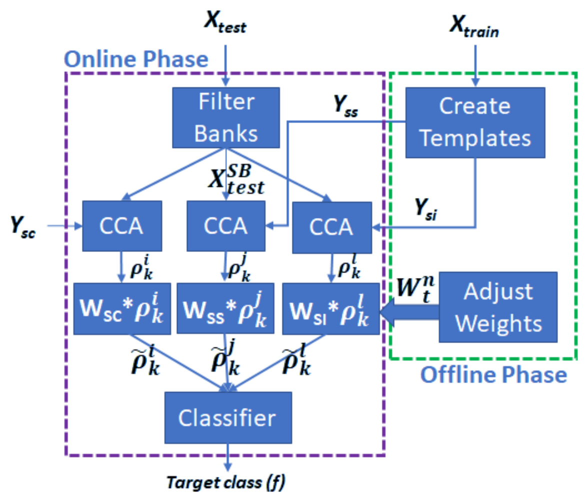

Fig. 1. Block Diagram of the proposed method: SCEF

Y_{SI}, using Equation. 4. Instead of the single optimized weight across subjects, we performed a grid search to identify the optimal weights per subject in each time window for each reference as W_{SC}, W_{SS} and W_{SI} respectively.

$$\bar{X} = \frac{1}{N}\sum_{t=1}^{N} X_t \tag{4}$$

Subjects' calibration references Y_{SS} and Y_{SI} provide useful Steady-State Responses characteristics pertained to intra-subject (inter-trials) and inter-subject information compared with subject-agnostic sine-cosine reference, Y_{SC} [23]. In online phase, test data X_{test} are firstly decomposed into n sub-bands resulting X_{test}^{SB}. Instead of only sine-cosine ideal reference, Y_{SC}, in FBCCA, the decomposed sub-bands X_{test}^{SB} are fed into three standard CCA operations with three references Y_{SC}, Y_{SS} and Y_{SI} respectively to compute correlation coefficients independently. Three correlation vectors ρ_k^{SC}, ρ_k^{SS} and ρ_k^{SI} where $k = [1, 2, \ldots, N_f]$ with respective weight vectors using Equation. 3 resulted individually weighted sum of squared correction vectors:namely, $\rho_{\hat{SC}k}$, $\rho_{\hat{SS}k}$ and $\rho_{\hat{SI}k}$ respectively. Subject-specific weight vector can be defined as $W_{SS} = [W_1^i, W_2^i, \ldots, W_{N_t-1}^i]$ where $i = [1, 2, \ldots, N_s]$ such that N_s is number of subjects. Subject-independent weight can be defined as $W_{SI} = [W_1, W_2, \ldots, W_{N_s-1}]$ including all trials from $N_s - 1$ subjects. The frequency d_f of target detection can be computed by finding the pairwise maximum from three independent weighted correlation coefficients from three references, $i = [SC, SS, SI]$ and stimulus frequencies $k = [f_1, f_2, \ldots, f_{N_f}]$ as shown in Equation. 5.

$$f_{SCEF} = \max_{(i,k)}(\hat{\rho}_k^{sc}, \hat{\rho}_k^{ss}, \hat{\rho}_k^{si}) \tag{5}$$

III. SSVEP Dataset and Evaluation Analysis

We use an open access benchmark data-set provided by Tsinghua University [24]. This data-set also was used in several other SSVEP studies to test performance among different methods [14], [15]. The data-set was from visual spelling task consisting of forty flickering targets by coding unique frequency and phase information for each stimulus. The stimuli are coded with frequencies range of 8-15.8 Hz with an interval of 0.2 Hz in combination with four distinct phases of $0, 0.5\pi, \pi and 1.5 pi$ radians. The details of stimulus layout and frequency-phase characteristics can be found in [24]. Thirty-five healthy subjects were recorded where each subject performs six blocks in a single session experiment. In each block, forty trials corresponding to each stimulus frequency were presented in a random order. The duration of each trial is 6 s, in which the first and last 0.5 s were used for visual cue and rest and 4s as SSVEP data. The provided EEG data are down-sampled at 250 Hz with 64 channels available in four-way tensors as $[N_c, N_s, N_f, N_t]$.

For creating references from calibration data, leave one trial out and leave one subject out approach is used in subject-specific, Y_{SS} and subject-independent Y_{SI} templates respectively. We used both accuracy and Information Transfer Rate (ITR) as performance metrics in comparison among methods. Firstly, we compared subject-specific weights with single weight per data-set on how much performance gains can be obtained. Then, we test the performance difference among FBCCA with three different references together with the proposed method, SCEF to understand the performance variability among different reference signals. FBCCA method with single CCA operation using ideal, subject-specific and subject-independent references can be defined as FBCCA (SC), FBCCA(SS) and FBCCA(SI) respectively. Finally, we compare the proposed method with state-of-the-art training method, Task Related Component Analysis (TRCA). TRCA requires subject-specific training that learns spatial filters to extract task-related components by maximizing the reproducibility of Steady-State Responses from multiple trials during the SSVEP task interval [25]. Nakanishi et al proposed TRCA spatial filtering by exploiting task-related components of Steady-State Responses by linear and weighted sum of multiple time courses that optimizes the maximum co-variance among trials [14].

We consider two ranges of time window from 4 s SSVEP data segment length: first is from 0.2 to 1 s with 0.1 s step (short time window) and, second is from 1.5 to 4 s with 0.5 time step (long time window). We only include time window up to 4 s although stimulus interval is 5 s. The reasons for this time window selection are in two folds: maximum perfect accuracy can mostly be reached at 4 s and, not much performance difference exists between time window higher than 4 s. In our offline analysis, we set the ranges of a and b from 0.0 to 2.0 and 1.0 to 1.0 with 0.25 step respectively to adjust subject-specific weights and compare its influences on performance gains. The values of N_c, N_h, N_t, N_f and N_s are set at nine channels(PO3, POz, PO4, O1, Oz, O2, PO7, PO8, Pz), five harmonics, six trials, forty frequencies and thirty-five subjects respectively. Similar to existing studies, Leave one trial out evaluation is used for computing TRCA performance [14]. For accuracy and ITR comparison, we used one-way repeated-measures analysis of variance (ANOVA) with Greenhouse–Geisser correction if spherical assumption of data is violated to test the significant differences among different methods. The statistical threshold for p-value is set at 0.001 and 0.05 in all the tests.

IV. Results and Discussion

We first test the effectiveness of different reference templates on target detection performance of SCEF comparing with single reference FBCCA methods with three reference types: FBCCA(SC), FBCCA(SS) and FBCCA(SI). We compared the accuracy of four methods in different time windows by using single weight coefficient $W_n = (1.25, 0.25)$ used in previous studies [14], [22] without subject-specific weights adjustment. This will allow us to compare how FBCCA performs differently with each reference type and SCEF method with subject-agnostic weight. Due to unique target discriminative characteristics of different reference types, SCEF outperforms all FBCCA variants using single weight in all time windows as shown in Figure. 2. Less similarity in target detection across trials in less than 1 s time windows cause SCEF achieving significantly better performance than FBCCA with any single reference type.

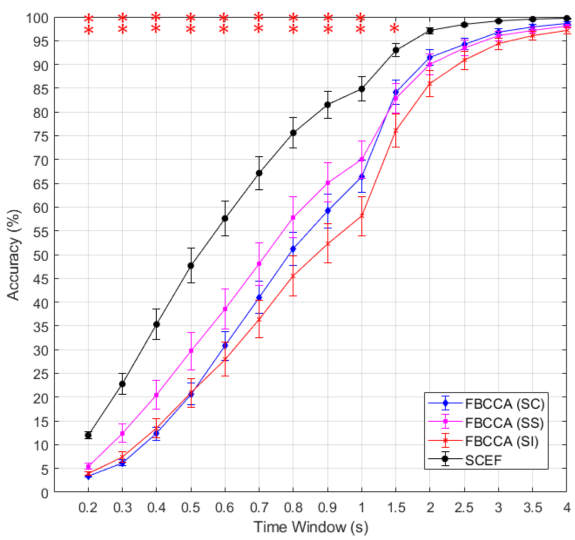

Fig. 2. Accuracy comparison among the proposed method, SCEF and FBCCA methods with three different references independently using the same weight coefficient of the data-set ($W_n = [1.25, 0.25]$) over all time windows.At specific time window, ** shows significant difference at p $<$ 0.001, * shows significant difference at p $<$ 0.05, otherwise ' ' for no statistical significant difference. Error bar shows standard errors.

One-way repeated measures of ANOVA results show highly significant difference among methods in time window from 0.2 to 1 s ($F[3, 99] = 25.48 \pm 13.39, p < 0.001$). But only significant difference at 1.5 s of longer time window range ($F[3, 99] = 3.22, p = 0.003$) and no significant difference in remaining time windows as shown in Table. I. This shows each reference template contribute independently and differently in target detection, especially in short time windows. These results highlight the important of using linear features extracted from three references independently in the proposed SCEF method.

Interestingly, subject-specific reference FBCCA(SS) only performs better than baseline FBCCA(SC) in time windows less than or equal to 1 s. This explains that subject-specific calibration can improve SSVEP detection performance in short time window. But there is no much advantage of subject-specific calibration in long time window as subject-agnostic ideal reference can reach accuracy higher than 90%. Because of the inter-subject differences in Steady-State Responses, subject-independent reference, FBCCA(SI) performs poorer than baseline FBCCA(SC) except time window less than 0.5 s. But all FBCCA and SCEF methods further improve the performance when different subject-specific weights in features fusion are used instead of optimal subject-agnostic weight for the dataset, $W_n = (1.25, 0.25)$. The following Figure. 3 shows accuracy comparison of the SCEF method between subject-specific weight and subject-agnostic weight in different time windows.

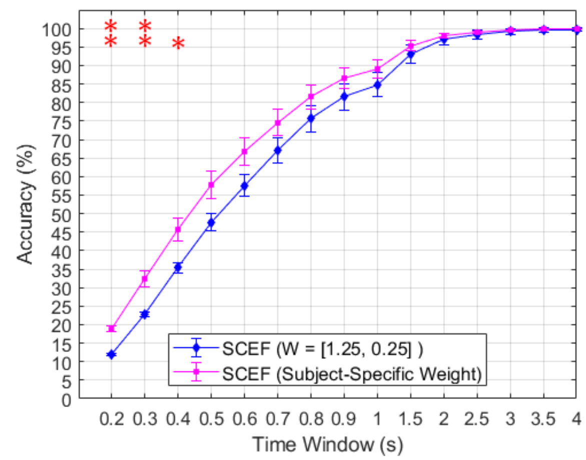

Fig. 3. Accuracy comparison of SCEF between subject-specific weights $W_{SCEF} = [W_{SC}, W_{SS}, W_{SI}]$ and single subject-agnostic weight $W_{SCEF} = (1.25, 0.25)$.At specific time window, ** shows significant difference at p $<$ 0.001, * shows significant difference at p $<$ 0.05, otherwise ' ' for no statistical significant difference.Error bar shows standard errors.

As expected, we can see almost similar high accuracy increments from 0.2 to 1 s time window but, smaller and different accuracy increments in 1.5 s onward. As performance in longer time window is reaching near perfect accuracy, our results did not attain much performance improvement. But the pairwise comparison results using Wilcoxon Rank-Sum test show only significant difference in three time windows of 0.2 - 0.3 s ($p < 0.001$) and 0.4 s ($p < 0.05$). This concludes that subject-specific weight improves the target classification performance of the proposed SCEF method consistently but differently across at time window as shown in Figure. 3. We further examine similar accuracy improvements in terms of difference single subject-agnostic weight and subject-specific weight in features fusion with SCEF and three FBCCA variants. Figure. 4 presents relative accuracy difference of each method between subject-agnostic and subject-specific weighted features fusion. Compared with FBCCA methods, SCEF method exhibits higher accuracy difference in time window less than 0.6s but lower accuracy difference in longer time window as shown in Figure. 4.

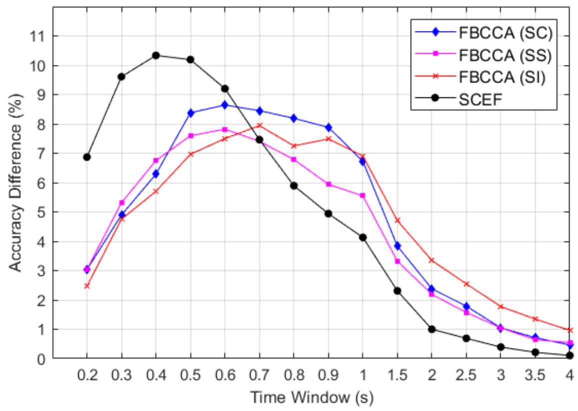

Fig. 4. Accuracy Difference between subject-agnostic weight (single weight per dataset) and subject-specific weight (weight per subject) among different methods.

We also observed that the mean accuracy differences of each method in both short and long time ranges between subject-agnostic and subject-specific weight in Figure. 4. The proposed SCEF method achieves high mean accuracy difference of 8.06 ± 2.06 in 0.2 to 0.9s time windows. But SCEF method only exhibits lower mean accuracy difference of 1.26 ± 1.47 % in 1 to 4 s time window. FBCCA methods with single reference exhibit similar performance improvement in both time windows. But FBCCA(SI) method has a bit higher mean accuracy differences of 3.08 ± 2.11 % compared with FBCCA(SC), 2.42 ± 2.22 % and FBCCA(SS), 2.12 ± 1.79 % in long time windows. All FBCCA variants have similar mean relative accuracy difference of 6.53 ± 1.83% in short time window.

Figure. 5 shows accuracy and ITR performance comparison of SCEF and three FBCCA variant methods using subject-specific weights. Statistical analysis using one-way repeated measures of ANOVA shows highly significant differences ($p < 0.001$) in time window 0.2 to 1 s and significant different ($p < 0.05$) at 1.5s in both accuracy and ITR. But there is no statistical significant difference among methods in long time window from 2 to 4s for both accuracy and ITR as shown in Table. I. Compared with accuracy of similar four methods in Figure. 2, all methods exhibit similar accuracy improvements with relatively higher accuracy as shown in Figure. 4 across time window. SCEF method achieves the highest ITR of 179 bits per minute at 0.9 s. But FBCCA(SC) and FBCCA(SS) achieve their maximum ITR of 128 and 137 bits per minute respectively at 1 s. The possible explanation on performance improvements by the proposed SCEF method compared with all variants of FBCCA in all time windows is for exploiting advantages in unique discriminant power of three references.

Although the performance of the proposed method is lower accuracy than basic TRCA in time window 0.2-0.3s as

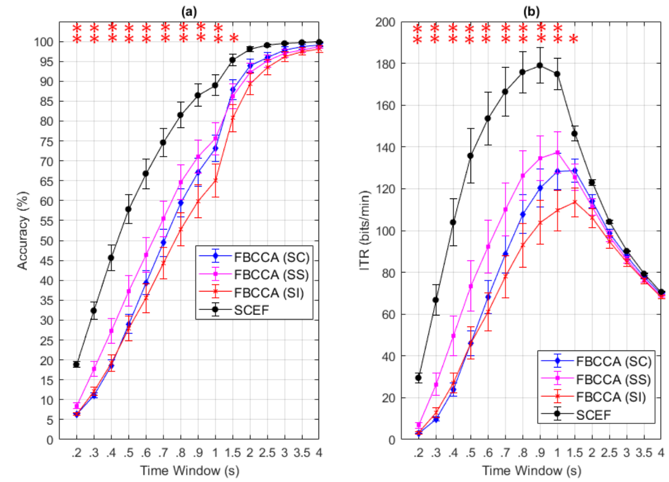

Fig. 5. Performance comparison among SCEF and FBCCA methods in different time windows (a) Mean Accuracy (b) Mean ITR. Error bar shows standard errors.At specific time window, ** shows significant difference at p < 0.001, * shows significant difference at p < 0.05, otherwise ' ' for no statistical significant difference.

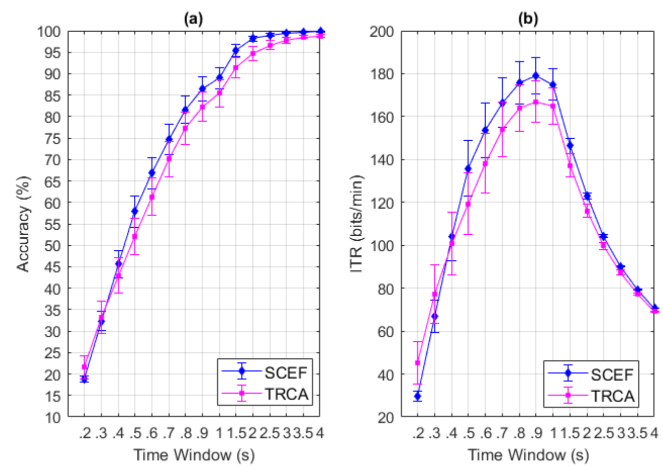

Fig. 6. Performance comparison between SCEF and TRCA methods in all time windows (a) Mean Accuracy (b) Mean ITR. Error bar shows standard errors.

shown in Figure. 6, there are no significant accuracy differences (F[1,33]=2.37, p=0.133) and (F[1,33]=0.73, p=0.399) at 0.2 and 0.3s respectively. Similarly, no significant difference between ITR of (F[1,33]=2.67, p=0.112) at 0.2 s and (F[1,33]=1.31, p=0.261) at 0.3s. But the proposed method shows consistently better performance in remaining time window from 0.4s to 4s. But there are no statistically significant differences between them in both accuracy and ITR as shown in Table II. The proposed method achieved 12 bits per minute higher than TRCA method at 0.9s where both methods have maximum ITR. These results highlight SCEF marginally perform better than TRCA in terms of 2.69 ± 2.32 and 5.51 ± 8.97 in overall accuracy and ITR averaged across all time windows.

We also validate the requirements of three unique reference

TABLE I
ONE-WAY REPEATED MEASURES OF ANOVA OF ACCURACY AND ITR RESULTS AMONG SCEF AND FBCCA VARIANTS FROM 2 TO 4S.

Parameters	Results	2 s	2.5 s	3 s	3.5 s	4 s
Accuracy	F[3,99]	0.93	0.55	0.22	0.29	0.27
	p-value	0.299	0.427	0.592	0.474	0.366
ITR	F[3,99]	1.82	1.16	0.45	0.57	0.5
	p-value	0.126	0.24	0.493	0.406	0.348

types used in the proposed method. This evaluation will shed lights on the substantial performance improvement of SCEF comparable to state-of-the-art method, TRCA, in short time window. We specify SCEF with only two references out of three available reference types (Y_{SC}, Y_{SS}, Y_{SI}) resulting three SCEF variants as shown in Figure 7. The one-way repeated measures of ANOVA results show high statistically significant difference in short time window from 0.2 s to 1 s ($p < 0.001$). From Figure 7, the similar accuracies among three SCEF variants with two reference types can be seen in all time windows. But SCEF (using all three reference types) outperform all SCEF variants (using only two reference types) especially in short time windows of 0.2 - 1 s. These results further validate that each reference even subject-independent can contribute to overall performance improvement of the proposed SCEF method.

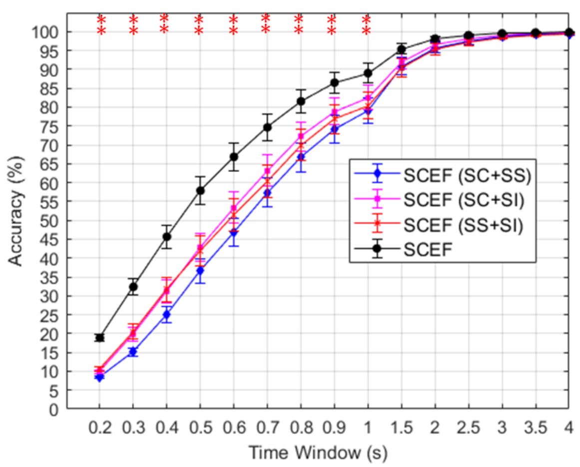

Fig. 7. Comparison of SCEF with two unique references combinations. Error bar shows standard errors. At specific time window, ** shows significant difference at p < 0.001, * shows significant difference at p < 0.05, otherwise ' ' for no statistical significant difference.

As we explained marginal performance improvement because of subject-specific weight, we identified the weight for each subject that can achieve higher mean accuracy as shown in Figure 8. But Some subjects have similar weight such as subject 12, 30 and 32 with $W_n = [1.25, 0]$ for each subject-specific weight adjustment. For each reference type, we need to identify separate subject-specific weights resulting to adjust three weights Y_{SC}, Y_{SS} and Y_{SI} as shown in Figure 1. As explained above, most of the performance improvements are mainly due to different discriminative decoding contributed by three reference types. We had not yet investigated whether significant performance differences exist among different weight pairs as we only grid search optimal weight for maximum mean accuracy per subject. In terms of subject-specific weight adjustment, we assumed that calibration data of individual subjects are not much difference across sessions and days in signal properties due to reliable and consistent state-state response from SSVEP paradigm [5]. As currently data-set only has single session data, it will be interesting to further test the performance using SSVEP data conducted in multiple sessions and days [26]. Such issue also applies to creating subject-specific reference Y_{SS} to determine whether in situ calibration requires and performance difference between pre-defined and just-in-time subject-specific reference. This is important as it is highly desirable calibration does not require performing as subject in every SSVEP session to avoid user's inconvenience and reduce overall experiment duration. With CCA approach, we assume SSVEP responses are linear in according to stimuli characteristics but SSVEP responses can be non-linearly originated [27]. Recently, linear correlation combined with non-linear temporal alignment shows improvement in accuracy without requiring calibration and subject-independently using 4-class SSVEP dataset [12], [26].

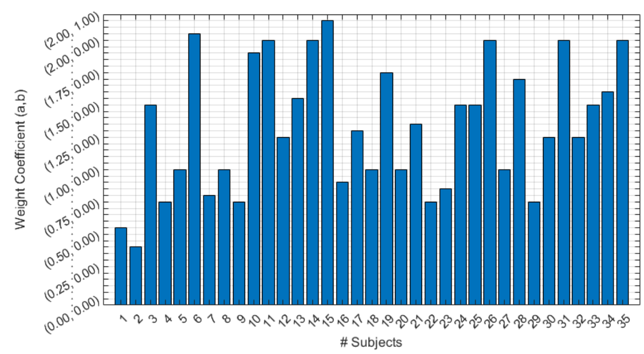

Fig. 8. Subject-specific Weight Assignment of W_{SS} for individual Subjects

We still need to validate whether the performance improvements by subject-specific weight vector can still apply to other data-set with different frequencies, number of targets, etc [17], [26]. We also highlight that subject-specific weighted feature fusion is less importance in accuracy improvements than subjects' calibrated references together with ideal sine-cosine reference as shown in Figure. I. The mutually exclusive frequency recognition of three references, less agreement in detection, allows notable performance improvements in short time windows. In terms of computational complexity, there is no additional computing intensive operation incurred as CCA operations with three references perform independently from EEG sub-bands inputs. Although SSVEP responses are reliable and highly consistent across trials and subjects, subject-independent and cross-subject frequency recognition are still not reliable though improvements shown through transfer learning approaches [11], [19]. Relying on CCA-based linear correlation features limits SSVEP decoding performance as SSVEP responses include both linear and non-linear from both

TABLE II
1-WAY ANOVA WITH REPEATED MEASURES BETWEEN SCEF AND TRCA.

Parameters	Time(s)	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	1.5	2	2.5	3	3.5	4
Accuracy	F[1,33]	2.37	0.73	0.34	0.59	0.04	0.01	0.33	0.10	0.005	0.0.36	0.61	1.21	0.85	1.55	2.70
	p-value	0.13	0.40	0.56	0.448	0.85	0.92	0.57	0.75	0.99	0.55	0.44	0.28	0.36	0.22	0.11
ITR	F[1,33]	2.67	1.31	0.70	0.65	0.04	0.03	0.56	0.07	1.11	1.52	2.35	1.40	2.62	0.29	4.02
	p-value	0.11	0.26	0.41	0.43	0.85	0.87	0.46	0.8	0.3	0.14	0.25	0.125	0.22	0.56	0.05

experimental and modeling study [27]. We will explore other oscillatory or temporal detection methods in replacing CCA operation to improve SSVEP features detection [12], [28]. Also, our evaluation is limited to baseline FBCCA and basic TRCA methods only as initial performance validation. Several improvements of TRCA were introduced by exploiting per-class calibration, multiple-stimulus calibration, etc [14], [15], [29]. In addition, We will further investigate on calibration requirements on templates creation and test the effectiveness of the SCEF method with different SSVEP data-sets.

V. Conclusion

To achieve a fast response SSVEP-based BCI, frequency recognition method must provide high accuracy with less performance variability in short time window less than 1 s. Several methods that exploit subject-specific calibration had shown accuracy improvements in short time windows. These methods with Filter-bank extensions enabled further accuracy improvements. So we optimized the FBCCA method by exploiting subjects' calibration templates, subject-specific weights adjustment with multiple reference templates to improve performance in short time window. Our empirical evaluation with forty-target SSVEP data-set shows that the proposed SCEF method achieved high statistically significant performance improvements compared with baseline FBCCA and its variants. Also, SCEF method exhibits overall better performance compared with basic TRCA method. Our analysis results highlighted that maximum performance gains can be obtained with independent operation and fusion of three reference templates. But we found that the contributions of subject-specific weights on performance improvements are less prominent than multiple reference templates. These improvements showed the potential performance gains for any method that uses CCA as spatial filtering and filter-bank as preprocessing step. Nevertheless, We will further test whether these optimization approaches can be applicable to other methods to improve performance. We are also planning to investigate whether performance gains can be seen in other SSVEP data-sets with different stimulus characteristics in near future.

Acknowledgment

This research was supported by Alibaba Group Holding Limited, DAMO Academy, Health-AI division. The program is the collaboration between Alibaba and Nanyang Technological University, Singapore.

References

[1] C. Brunner, N. Birbaumer, B. Blankertz, C. Guger, A. Kübler, D. Mattia, J. d. R. Millán, F. Miralles, A. Nijholt, E. Opisso *et al.*, "BNCI horizon 2020: towards a roadmap for the BCI community," *Brain-computer interfaces*, vol. 2, no. 1, pp. 1–10, 2015.

[2] M. J. Wieser, V. Miskovic, and A. Keil, "Steady-state visual evoked potentials as a research tool in social affective neuroscience," *Psychophysiology*, vol. 53, no. 12, pp. 1763–1775, 2016. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/psyp.12768

[3] X. Chen, B. Zhao, Y. Wang, S. Xu, and X. Gao, "Control of a 7-dof robotic arm system with an ssvep-based bci," *International Journal of Neural Systems*, vol. 28, no. 08, p. 1850018, 2018, pMID: 29768990. [Online]. Available: https://doi.org/10.1142/S0129065718500181

[4] X. Chen, Z. Chen, S. Gao, and X. Gao, "A high-ITR SSVEP-based BCI speller," *Brain-Comp.Interfaces*, vol. 1, no. 3-4, pp. 181–191, 2014.

[5] A. M. Norcia, L. G. Appelbaum, J. M. Ales, B. R. Cottereau, and B. Rossion, "The steady-state visual evoked potential in vision research: A review," *Journal of Vision*, vol. 15, no. 6, pp. 4–4, 05 2015.

[6] M. Nakanishi, Y.-T. Wang, T.-P. Jung, J. K. Zao, Y.-Y. Chien, A. Diniz-Filho, F. B. Daga, Y.-P. Lin, Y. Wang, and F. A. Medeiros, "Detecting glaucoma with a portable brain-computer interface for objective assessment of visual function loss," *JAMA ophthalmology*, vol. 135, no. 6, pp. 550–557, 2017.

[7] A. A. Phyo Wai, Z. Goh, S. D. Foo, and C. Guan, "A study of ssvep responses in case of overt and covert visual attention with different view angles," in *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Oct 2018, pp. 3847–3852.

[8] D. Won, H. H. Zhang, C. Guan, and S. Lee, "A bci speller based on ssvep using high frequency stimuli design," in *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Oct 2014, pp. 1068–1071.

[9] Z. Zhang, C. Wang, K. K. Ang, A. A. P. Wai, and C. G. Nanyang, "Spectrum and phase adaptive cca for ssvep-based brain computer interface," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, July 2018, pp. 311–314.

[10] Y. S. Zhang, E. Yin, F. Li, Y. Zhang, T. Tanaka, Q. Zhao, Y. Cui, P. Xu, D. Yao, and D. Guo, "Two-stage frequency recognition method based on correlated component analysis for SSVEP-based BCI," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 7, pp. 1314–1323, 2018.

[11] R. Zerafa, T. Camilleri, O. Falzon, and K. P. Camilleri, "To train or not to train? a survey on training of feature extraction methods for SSVEP-based BCIs," *J. Neural Eng.*, vol. 15, no. 5, p. 051001, 2018.

[12] A. A. P. Wai, M.-H. Lee, S.-W. Lee, and C. Guan, "Improving the performance of SSVEP BCI with short response time by temporal alignments enhanced CCA," in *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*. IEEE, 2019, pp. 155–158.

[13] Z. Lin, C. Zhang, W. Wu, and X. Gao, "Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 6, pp. 1172–1176, 2007.

[14] M. Nakanishi, Y. Wang, X. Chen, Y. T. Wang, X. Gao, and T. P. Jung, "Enhancing detection of SSVEPs for a high-speed brain speller using task-related component analysis," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 1, pp. 104–112, Jan 2018.

[15] Y. S. Zhang, E. Yin, F. Li, Y. Zhang, D. Guo, D. Yao, and P. Xu, "Hierarchical feature fusion framework for frequency recognition in SSVEP-based BCIs," *Neural Networks*, vol. 119, pp. 1–9, 2019.

[16] X. Chen, Y. Wang, M. Nakanishi, X. Gao, T.-P. Jung, and S. Gao, "High-speed spelling with a noninvasive brain-computer interface," *Proc.Natl. Acad. Sci. USA.*, vol. 112, no. 44, pp. E6058–E6067, 2015.

[17] M. Nakanishi, Y. Wang, Y.-T. Wang, and T.-P. Jung, "A comparison study of canonical correlation analysis based methods for detecting steady-state visual evoked potentials," *PLoS ONE*, vol. 10, no. 10, p. e0140703, 2015.

[18] N. R. Waytowich, V. Lawhern, J. O. Garcia, J. Cummings, J. Faller, P. Sajda, and J. M. Vettel, "Compact convolutional neural networks for classification of asynchronous steady-state visual evoked potentials," *arXiv preprint arXiv:1803.04566*, 2018.

[19] K. Chiang, C. Wei, M. Nakanishi, and T. Jung, "Cross-subject transfer learning improves the practicality of real-world applications of brain-computer interfaces," in *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*, March 2019, pp. 424–427.

[20] A. M. Azab, "A review on transfer learning approaches in brain–computer interface: Signal processing and machine learning for brain-machine interfaces," pp. 81–101, 2018.

[21] X. Chen, Y. Wang, S. Gao, T.-P. Jung, and X. Gao, "Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain-computer interface," *J. Neural Eng.*, vol. 12, no. 4, p. 046008, 2015.

[22] G. R. Kiran Kumar and M. Ramasubba Reddy, "Filter bank extensions for subject non-specific ssvep based bcis," in *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*, March 2019, pp. 627–630.

[23] P. Yuan, X. Chen, Y. Wang, X. Gao, and S. Gao, "Enhancing performances of SSVEP-based brain-computer interfaces via exploiting inter-subject information," *J. Neural Eng.*, vol. 12, no. 4, p. 046006, 2015.

[24] Y. Wang, X. Chen, X. Gao, and S. Gao, "A benchmark dataset for SSVEP-based brain-computer interfaces," *IEEE Trans.Neural Syst. Rehabil.Eng.*, vol. 25, no. 10, pp. 1746–1752, Oct 2017.

[25] H. Tanaka, T. Katura, and H. Sato, "Task-related component analysis for functional neuroimaging and application to near-infrared spectroscopy data," *NeuroImage*, vol. 64, no. Supplement C, pp. 308 – 327, 2013.

[26] M.-H. Lee, O.-Y. Kwon, Y.-J. Kim, H.-K. Kim, Y.-E. Lee, J. Williamson, S. Fazli, and S.-W. Lee, "EEG dataset and OpenBMI toolbox for three BCI paradigms: an investigation into BCI illiteracy," *GigaScience*, vol. 8, no. 5, 01 2019, giz002. [Online]. Available: https://doi.org/10.1093/gigascience/giz002

[27] M. Labecki, R. Kus, A. Brzozowska, T. Stacewicz, B. S. Bhattacharya, and P. Suffczynski, "Nonlinear origin of ssvep spectra—a combined experimental and modeling study," *Frontiers in computational neuroscience*, vol. 10, p. 129, 2016.

[28] B. Wittevrongel and M. M. Van Hulle, "Frequency- and phase encoded ssvep using spatiotemporal beamforming," *PLOS ONE*, vol. 11, no. 8, pp. 1–18, 08 2016. [Online]. Available: https://doi.org/10.1371/journal.pone.0159988

[29] C. M. Wong, F. Wan, B. Wang, Z. Wang, W. Nan, K. F. Lao, P. U. Mak, M. I. Vai, and A. Rosa, "Learning across multi-stimulus enhances target recognition methods in SSVEP-based BCIs," *Journal of Neural Engineering*, vol. 17, no. 1, p. 016026, jan 2020. [Online]. Available: https://doi.org/10.1088/1741-2552/ab2373