Enhancing the classification accuracy of Steady-State Visual Evoked Potential-based Brain-Computer Interface using Component Synchrony Measure

Kian B. Ng, Ross Cunnington
The University of Queensland
Queensland Brain Institute
Brisbane, Queensland 4072, Australia
Email: k.ng1@uq.edu.au

Andrew P. Bradley
The University of Queensland
School of Information Technology and Electrical Engineering
Brisbane, Queensland 4072, Australia

Abstract—Steady-State Visual Evoked Potential-based (SSVEP) Brain-Computer Interface (BCI) shows great potential as a viable BCI due to its ease of implementation and speed. However, the majority of the SSVEP-BCI implementations use only features from the Power Spectral Density (PSD) despite the fact that upon transforming the signals to the Fourier domain, both the phase and amplitude components are available. In this study we extract the phase response and compute the phase variance as a measure of phase synchrony. This phase synchrony method is called Component Synchrony Measure (CSM). Our results indicate that by including the CSM as a feature, the SSVEP-BCI classification accuracy is significantly enhanced. This further establishes the use of both amplitude and phase information for obtaining good classification accuracy in SSVEP-BCI.

Keywords - component synchrony measure; visual evoked potential; brain-computer interface; phase variance

I. INTRODUCTION

In Steady-State Visual Evoked Potential-based Brain-Computer Interface (SSVEP-BCI), several frequency tagged visual stimuli are presented to user. These visual stimuli represent the on-screen options available to the user to perform a specific task. When the user focuses and fixates on a target stimulus, the VEP responses captured at the occipital cortex exhibit the same frequency and phase as that of the attended visual stimulus [1]. Therefore one way of labeling the BCI data in subsequent analysis is to transform the temporal signal to the frequency domain using Fast Fourier Transform (FFT) and consider the frequency and phase spectra as features in the classification process.

According to the survey conducted by Bashashati et al. [2], due to the ease of implementation, computational speed, and direct interpretation of results, most VEP-BCI designs predominantly use the power spectral density (PSD) as input features. The PSD represents the power per hertz of the signal. It measures the amplitudes of the FFT along the frequency spectrum. A dominant amplitude corresponds to the attended visual stimulus at the specific frequency. While the FFT components also include the phase information, the use of phase as a feature in BCI classification is not straightforward. To correctly measure the phase, the BCI needs precise and stable time-triggers from the visual simulator and recorder so that phase differences or consistencies can be reliably assessed. Furthermore, phase responses from SSVEP cannot be reliably determined using the short data segments often employed in real-time classification due to its sensitivity towards noise [3]. Hence, methods such as the use of a high-resolution timer in generating visual stimuli and increased window segment size are often employed in analyzing the phase response.

In recent studies, reports have suggested that the use of phase information can aid in realizing more visual targets under limited frequencies [4], [5] and improve target identification accuracy [6], [7]. Given that the phase and amplitude are independent, using both the amplitude and phase responses may enhance the performance of the BCI classification. In this paper, we investigate the use of phase to enhance SSVEP-BCI classification accuracy. We propose to use a method called Component Synchrony Measure (CSM) to evaluate the phase consistency as a feature in BCI classification. CSM estimates the phase variance of a given component frequency over a number of averaged responses. Together with Welch's method of estimating PSD [8], we are able to compare SSVEP-BCI classification accuracy. Our investigation consists of three parts: the first part is to use PSD feature alone, the second is to use CSM feature alone, and the third is to use both PSD and CSM features together. This will elucidate the effect of CSM in enhancing SSVEP-BCI classification accuracy. Our classification method for SSVEP-BCI is based on Support Vector Machine (SVM) [9].

In Section II of this paper, we will provide a brief description of CSM and the features we extracted to input to the SVM. In Section III, we will describe the experimental setup that allows us to study the classification accuracy under varying stimuli parameter conditions so that effect of PSD and CSM can be observed. We present the results in Section IV and conclude the paper in Section V with a discussion.
II. METHODS

A. Component Synchrony Measure

Component Synchrony Measure (CSM) estimates the degree of synchrony for a given Fourier component, \( m \), by taking its one complement to the phase variance [10]. The phase variance, according to Mardia [11], can be defined as:

\[
\varphi \{ \phi(m) \} = 1 - \left[ \frac{1}{n} \sum_{i=1}^{n} \cos \phi_i(m) \right]^2 - \left[ \frac{1}{n} \sum_{i=1}^{n} \sin \phi_i(m) \right]^2
\]

(1)

where \( \phi_i(m) \) is the phase of the \( m \)th Fourier component, and \( n \) is the number of averaged waveforms. Hence CSM can be computed as:

\[
CSM(m) = 1 - \varphi \{ \phi(m) \}
\]

(2)

To compute the CSM, we first take each input epoch and divide the data into overlapping window segments. Next, we transform the signal using the FFT. If the FFT length is greater than the window segment length, the signals are zero padded. The phase of each component frequency is then calculated and the variance is computed as in (1). The results of the CSM from (2) range from 0 to 1. When the phase is synchronized within the epoch, the phase variance is close to zero and thus CSM yields a measure close to 1. Conversely, when the phase is not synchronized, CSM returns a value close to zero. Fig. 1 shows an example of the CSM values we obtained in our experiment where a dominant CSM value was seen at a stimulus frequency at 11.3 Hz. In our implementation, we integrated our CSM development with the PSD code available from Octave for MATLAB [12]. The PSD from Octave uses Welch's estimate and involves a similar approach to our CSM method where the FFT responses were utilized to obtain the spectral estimates. This integration eases our effort in feeding the PSD and/or CSM as features to our SVM classifier.

B. Feature Extraction

We derived two features each from the PSD and CSM as shown in Table 1. In the Table 1, \( P_{\text{attended}} \) denotes the PSD obtained at the attended stimulus frequency, \( mP_{\text{untag}} \) denotes the mean PSD at the untagged frequencies, and \( P_{\text{max}} \) denotes PSD of the maximum peak at any of the other tagged frequencies. Similarly, \( C_{\text{attended}} \) denotes the CSM obtained at the attended frequency, \( mC_{\text{untag}} \) denotes the mean CSM at the untagged frequencies, and \( C_{\text{max}} \) denotes CSM of the maximum peak at any of the other tagged frequencies. Hence the \( mP_{\text{untag}} \) and \( mC_{\text{untag}} \) give an indication of the level of the background noise. In this way, feature 1 estimates the signal to noise ratio while feature 2 indicates the strength of the PSD and CSM of the attended stimulus compared to the strongest of the unattended stimuli. These features are fed into our SVM for classification.

| Features Extracted from the PSD and CSM. |
|---|---|
| **Features** | **PSD** | **CSM** |
| 1 | \( P_{\text{attended}} \) | \( C_{\text{attended}} \) |
| 2 | \( mP_{\text{untag}} \) | \( mC_{\text{untag}} \) |
| 3 | \( P_{\text{max}} \) | \( C_{\text{max}} \) |

C. Classification

The SVM was originally designed for two-class classification. Given that there are usually more than two tagged visual stimuli on screen (in our experiment, there were six) in SSVEP-BCI, it is necessary to perform multiclass SVM. With the help of the open source machine learning library, dlib [13], we implemented a multiclass support vector machine (MSVM). Our MSVM adopts the one-versus-one approach as shown in (3):

\[
\min_{w, \xi} \frac{1}{2} (w^j)^T w^j + C \sum \xi_i (\phi(x_i))^T w^i
\]

\[
\text{subject to } (w^i)^T \phi(x_i) + b^i \geq 1 - \xi_i, \text{if } y_i = i
\]

\[
(w^j)^T \phi(x_i) + b^j \leq -1 + \xi_i, \text{if } y_i = j
\]

\[
\xi_i \geq 0
\]

where the training data \( x_i \) from \( i \)th and \( j \)th classes are mapped to a higher dimensional space by the function and \( C \) is the penalty parameter. The kernel function employed here is the radial-basis function (RBF) as suggested in by Chih-Wei Hsu [14]. The RBF is defined as follow:

\[
K(x, y) = e^{-\gamma \| x - y \|^2}
\]

where \( \gamma \) is the kernel parameter.

The penalty parameter \( C \) controls the tradeoff between the training error and model complexity. Together with the kernel parameter \( \gamma \), the \( (C, \gamma) \) values will influence the accuracy of
the classification. For a detailed discussion on SVM model parameters, refer to [15–17]. Here, we performed parameter estimation using a leave-one-out cross validation procedure. Leave-one-out is the recommended strategy when the dataset is of limited size [18]. The decision on the values of the penalty parameter \( C \) and the kernel parameter \( \gamma \) was based on the accuracy results from the grid-search [14].

III. EXPERIMENTAL SETUP

A. Subjects

A total of seven healthy subjects, one male and six females, participated in the experiment. All were right handed and between the ages of 21 and 28. All had normal or corrected-to-normal vision. None of them had any history of neurological disorder. Written informed consent was obtained from each participant prior to the beginning of the experiment.

B. EEG Acquisition

The experiment was conducted in a Faraday-caged room to reduce any electromagnetic interference from external sources. Continuous EEG was acquired using a 64-channel Biosemi ActiveTwo system. Two additional electrodes, CMS (Common Mode Sense; active electrode) and DRL (Driven Right Leg; passive electrode) were used to compose a feedback loop for amplifier reference. The electrodes used were placed according to the international 10-20 system using an appropriate EEG cap that fitted the head size of the individual subject. The EEG data were sampled at a rate of 2048 Hz to 24 bit precision.

C. Stimuli

In the survey conducted by Zhu et al. [19], the majority of SSVEP-BCI implementations used frequencies in the alpha and beta band. Even though higher band frequency (> 30 Hz) may cause less fatigue, it is difficult to implement due to hardware limitations [20]. As for lower frequency (< 10 Hz), they appear to be a major contaminant of the evoked response [21]. Thus, we chose to restrict our tagging frequency bands to between high alpha and beta band. Furthermore, to ensure reliable frequency and phase measurements, we developed an in-house presentation software to create and display the stimuli. The software made use of the high-resolution hardware timer available through the PC system library. The method of frequency display was based on the frame-based encoding method proposed by Y. Wang et al [22]. The method combines sequences of two frequencies into a single sequence to approximate a target frequency with varying number of frames in each cycle. The frequencies we implemented were 11.3, 13.6, 16.1, 18.3, 21.4 and 23.3 Hz. Their consistencies were verified using a photometer.

The stimuli were hexagons flashing in black and white displayed on a mid gray background using a LED monitor (Samsung 23” XL2370) with a response time of 2 milliseconds. The luminance of the white hexagon was 165 cd/m² whereas luminance of the black hexagon was 0.70 cd/m². The primary advantage of hexagonal stimuli over traditional square shaped stimuli is that the distance between the centre of each hexagon and centres of all adjacent hexagons can be made constant.

This ensures a consistent measurement of inter-stimulus distance across all stimuli.

The stimuli were each set to a constant size that subtended a visual angle of 2° from edge to edge. There were two main varying parameters, proximity and count. The proximity of the stimulus refers to the spatial distance between the attended stimulus and the unattended stimuli. The unattended stimuli were always placed concentrically around the attended stimulus. The distances, as measured in visual angles, were 2°, 3°, 5°, and 7°. The count factor refers to the number of competing stimuli shown simultaneously on the screen. There could be two, three, or four competing stimuli shown together with the central stimulus. Each of the six tagged frequencies would in turn be set as the central frequency. The arrangement ensured that all frequencies were equally distributed. Table II summarizes all the possible combinations of conditions. By varying the proximity and count, we were able to investigate the effect of the input features (PSD or/and CSM) have on the SSVEP-BCI classification accuracy.

<table>
<thead>
<tr>
<th>Proximity</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td>2°</td>
<td>(2°, 2)</td>
</tr>
<tr>
<td>3°</td>
<td>(3°, 2)</td>
</tr>
<tr>
<td>5°</td>
<td>(5°, 2)</td>
</tr>
<tr>
<td>7°</td>
<td>(7°, 2)</td>
</tr>
</tbody>
</table>

D. Experiment Paradigm

Throughout the experiment, subjects were seated 70 cm in front of the monitor. The monitor’s full screen subtended a visual view angle of 40°. Subjects were instructed to fixate on a central crosshair and focus their attention only on the flashing hexagon that appeared in the centre ignoring all other stimuli. The crosshair always appeared in the centre, therefore the attended stimulus was always centrally viewed.

For each trial, the flashing stimuli appeared for 8 s duration followed by 8 s with just the fixation crosshair. They were termed target and non-target epoch respectively. The subjects were asked to maintain their central fixation for 8 s during the target epoch. While VEPs can be captured in a much shorter time span, the 8 s duration gives ample resolution to avoid the time-frequency uncertainty [23] in subsequent analysis.

Across trials, the number of unattended stimuli (0, 2, 3, or 4) and the spatial separation of stimuli (2°, 3°, 5° and 7°) were varied. The tagged frequency of the central stimulus was also varied and tested once with each of the six different tagging frequencies. Fig. 2 shows an example of the experiment sequence using three competing stimuli.

To ensure that the subjects maintained fixation and attention to the central stimulus, they were told to note the frequency difference between each successive central stimulus.
If the frequency was perceived as faster than the previous, the subject would press the 'z' key, if slower, press the 'enter' key. In addition, to avoid frequency adaptation, the order in which the frequency appeared was varied randomly across different trials.

**E. Data Processing**

The EEG data obtained was stored locally in Biosemi’s 24-bit bdf format and processed offline. The data were first digitally filtered through a 50Hz notch filter and then a 4th order Butterworth band-pass filter with cut-off frequencies at 2 and 48 Hz. This frequency range allowed us to consider both the fundamental and second harmonic frequencies in subsequent analysis [24]. However, in our subsequent observation, the second harmonics for frequencies F5 (21.4 Hz) and F6 (23.3 Hz), that is 42.8 and 46.6 Hz, were absent. This observation is in line with the report by [25] where not all second harmonics of high frequencies are detectable. Hence, to avoid biasing the feature vector, only the fundamental harmonics were used.

Next, the common average reference (CAR) procedure was used to re-reference all electrodes. CAR is an averaging process, by which only common signals (including noise) from all electrodes remain in the reference signal; isolated signals do not remain unless they are large and dominating. Background noise that is uncorrelated and random with zero mean is also removed through this re-referencing procedure.

Artifacts arising from the electrooculographic (EOG) activity such as eye movements and eye blinks and electromyographic (EMG) activity from scalp and facial muscles are embedded in EEG signals. These artifacts are generally seen as high-amplitude patterns with large disturbances in the brain signals. In our analysis, they were detected using the outlier detection toolbox from MATLAB and visually inspected. The contaminated epochs were subsequently removed in our analysis. In addition, there were 64 channels of data available for every recorded EEG signal. However, the actual amount of data relevant to our analysis was anticipated to be substantially smaller since only signals at the occipital region would exhibit the SSVEP characteristics [1]. Hence, to help reduce data dimensionality, we have chosen to focus our analysis using only the occipital sites, i.e. O1, O2, and Oz.

Fig. 3 shows the entire data pipeline in our analysis. The parameters we used in obtaining the features were: Hamming window function of sample length 2048 (1 s), 1024 (50%) samples overlap, and 8192 samples of the FFT. As the FFT length was greater than the window segment length, the signals were zero padded. This zero padding helped to improve the accuracy of the measured amplitudes, at intervals of 0.25Hz. Given that the six tagging frequencies were not multiples of the frequency interval (i.e. 0.25 Hz), when deciding the PSD and CSM of the attended frequency, the closest pair of frequencies to the target frequency was taken into account (e.g. for 11.3 Hz, frequencies at 11.25 and 11.5 Hz were considered). The highest value between the pair was selected for both magnitude and phase analysis.

After extracting the PSD and CSM features as shown in Table 1, we fed the data to the MSVM and measured the classification accuracy. Given the three occipital sites, the input would be considered as correctly labelled as long as any two of the channels (Oz/O1/O2) successfully classified it. The outcome of the predicted stimulus frequency and the actual stimulus frequency were then used to present how the classification accuracy was influenced by the proximity and count. In the beginning, only PSD features were used. Next, only CSM features were used. Finally, both PSD and CSM features were used to study the classification accuracy.

**IV. RESULTS**

The outcome of the predicted stimulus frequency and the actual stimulus frequency were gathered into a confusion matrix. The confusion matrix provides an indication of the SSVEP-BCI classification accuracy, as a percentage, under varying stimulus conditions. An example of the confusion matrix for subject 1, using the PSD features only, under the condition of proximity of 7° with count at 2 is shown in Table III.
TABLE III. **CONFUSION MATRIX FOR SUBJECT 1, USING PSD FEATURES ONLY.** THE 'A' REPRESENTS ACTUAL TARGET FREQUENCY AND THE 'P' REPRESENTS THE PREDICTED FREQUENCY RESULT.

<table>
<thead>
<tr>
<th>A/P</th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
<th>F4</th>
<th>F5</th>
<th>F6</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>97.1</td>
<td>0</td>
<td>2.9</td>
<td>2.9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>F2</td>
<td>0</td>
<td>82.9</td>
<td>5.7</td>
<td>0</td>
<td>2.9</td>
<td>5.7</td>
</tr>
<tr>
<td>F3</td>
<td>0</td>
<td>0</td>
<td>97.1</td>
<td>0</td>
<td>0</td>
<td>2.9</td>
</tr>
<tr>
<td>F4</td>
<td>0</td>
<td>5.7</td>
<td>91.4</td>
<td>0</td>
<td>2.9</td>
<td>0</td>
</tr>
<tr>
<td>F5</td>
<td>5.7</td>
<td>0</td>
<td>0</td>
<td>88.6</td>
<td>5.7</td>
<td>0</td>
</tr>
<tr>
<td>F6</td>
<td>5.7</td>
<td>5.7</td>
<td>0</td>
<td>5.7</td>
<td>80</td>
<td>0</td>
</tr>
</tbody>
</table>

**Proximity = 7°, Count = 2**

Fig. 4 summarizes the results of the mean classification accuracy from all seven subjects under the varying stimuli parameter conditions (as in Table II), using firstly the PSD features alone, then the CSM features alone, and finally the PSD and CSM features together. From Fig. 4, it can be observed that at each of the stimulus parameter conditions, there appears to be an enhancement of the classification accuracy when both PSD and CSM features were used. However, between using PSD features alone and CSM features alone, it does not appear that either one is better. To validate our observations, we performed the Analysis of Variance (ANOVA) test.

A one-way ANOVA was carried out to examine the null hypothesis that the features we use (PSD alone, or CSM alone, or PSD+CSM together) make no difference to the classification accuracy. Results indicated significant differences in the classification accuracy from the different features: $F(2,429) = 33.29, p < 0.05$. Tukey post-hoc test was further performed to isolate the significant difference. The test revealed that significant differences ($p < 0.05$) were found between using PSD features alone and PSD+CSM features; CSM features alone and PSD+CSM features.

No study was conducted to investigate the effect of the stimulus parameters (e.g. the effect of stimulus proximity to classification accuracy) as this has been reported in a separate paper by Ng et al. [27].

**V. DISCUSSION**

Based on the report by Joon Kim, Y. et al [28], attention induces synchronization of the neural responses in SSVEP. Thus, it follows that that when a user fixates and focuses on a target frequency in an SSVEP-BCI, the neurons in the visual cortex will fire in synchrony, resulting in enhanced amplitude response and a phase that is distributed over a narrow range. Conversely, a widely distributed phase would indicate a more stochastic signal, which often manifests in weaker SSVEP amplitudes. However, in the case when there is substantial noise, the firing of the neurons and the SSVEP observed may become loosely coupled. For such instances, the irregularity of the amplitudes may mask this phase relationship. Techniques relying only on the signal power may thus become less sensitive in the detection of the features [29]. Hence, measuring...
the phase synchrony provides an additional dimension to improve our SSVEP-BCI classification performance. This study has elucidated the effect of phase, in the form of Component Synchrony Measure, in enhancing the classification accuracy of SSVEP-BCI.

For our future work, to demonstrate the efficacy of this method (PSD with CSM), we aim to perform a comparison study with other commonly used classification algorithms such as Canonical Correlation Analysis (CCA) [30] and Linear Discriminant Analysis (LDA).

ACKNOWLEDGMENT
The authors thank Katharine Baker for her assistance, especially at the initial stage of the experiments, and the subjects who participated in the EEG recording sessions.

REFERENCES