Phase synchrony analysis for SSVEP-based BCIs

Danhua Zhu1,2,3, Gary Garcia Molina2, Vojkan Mihajlovíc2, Ronald M. Aarts1,2
3. College of Biomedical Engineering and Instrument Science, Zhejiang University, Hangzhou, China.
E-mail: [zhu.danhua, gary.garcia, vojkan.mihajlovic, ronald.m.aarts]@philips.com

Abstract—Brain-computer interfaces (BCI) based on Steady State Visual Evoked Potential (SSVEP) can provide higher throughput than other BCI modalities. For the sake of safety and comfort, the frequencies of the stimulus should be higher than 30 Hz. However, only a limited number of these frequencies can elicit SSVEPs that are strong enough for BCI purposes. In order to increase the number of available stimuli, the SSVEP phase can be taken into account. In this study, we used phase synchrony analysis to extract the phase difference between SSVEP and stimuli as a feature to identify a subject’s intention. This analysis can mitigate the adverse effect brought by the phase deviation that may occur in the stimuli. Furthermore, the classification accuracy when using a single lead signal (Oz-Cz) is compared to a spatial filtered signal. The result shows that the phase synchrony analysis can effectively extract the phase difference and that spatial filtering can significantly increase the classification accuracy.

Keywords—brain computer interface (BCI); steady state visual evoked potential (SSVEP); Hilbert transform; phase synchrony; spatial filtering.

I. INTRODUCTION

The steady state visual evoked potential (SSVEP) refers to the response of the cerebral cortex to a repetitive visual stimulus (RVS) oscillating at a constant frequency and can be characterized by peaks at the fundamental frequency and its harmonics in the power spectral density (PSD) of EEG signals. The SSVEP is an effective electrophysiological source that can be used as an input of a brain computer interface (BCI). Previous studies have demonstrated that it is possible to detect to which one of a set of stimuli a subject is paying attention by examining the amplitudes at stimulus frequencies in the power spectral estimates of simultaneously recorded EEG signals [1]-[3].

Most current SSVEP-based BCIs use frequencies between 4 Hz and 30 Hz. The stimuli oscillating at these frequencies have a higher probability of causing visual fatigue and risk of epileptic seizure compared to ones oscillating at frequencies higher than 30 Hz [4]. These higher frequencies are therefore preferable from the viewpoint of safety and comfort. However, only a limited number of frequencies above 30 Hz can elicit a sufficiently strong SSVEP for BCI purposes (see Section IV). Thus, if only one frequency per target is used, the number of choices in a BCI is limited.

One way to tackle this problem is combining two frequencies to drive a visual stimulus [5], [6]; another way is using the same frequency but different phases [7], [8]. Compared to the former method, the latter completely maintains the periodicity of a stimulus. In addition, these studies show that using phase as a feature is feasible and promising, because SSVEP is phase-locked with the stimulus, i.e. the phase difference of a stimulus and the elicited SSVEP fluctuates around a certain value.

SSVEP phase can be obtained using the Discrete Fourier Transform (DFT) [7], [9] or the Short Time Fourier Transform (STFT) [8]. These methods require a relatively long signal whose duration is a multiple of the stimulus period. In addition, current research only considers the absolute SSVEP phase, which requires the synchronization of presenting stimuli and recording EEG signals between training and operation sessions.

In this study, we incorporated the stimulus signal recorded by a photodiode and used the Hilbert transform to extract the phase difference between SSVEP and the stimulus signal. In our experience, this phase difference is dependent solely on the subject and the frequency. Because the subtraction of the two phases from the SSVEP and the stimulus signal cancels the time variable out, the synchronization of presenting stimuli and recording EEG signals between sessions is not necessary. An additional advantage of using the phase difference as a feature is that it can avoid the change in SSVEP phase distribution caused by the possible phase deviation of stimuli.

Another aspect addressed in this study is the performance increase that can be gained from combining several electrode channels into one signal using a spatial filter, compared to using a single lead signal (Oz-Cz). The results showed that spatial filtering can significantly increase the separability of the phase difference.

This paper is organized as follows. In Section II, the method to enhance SSVEP and extract the phase difference between SSVEP and the stimulus signal is presented. Section III addresses the experimental setup. Section IV and Section V explain and discuss the results, respectively. The conclusions and prospective future research are presented in Section VI.

II. METHODS

In this section, we summarize the method for spatial filtering based on [10], [11] and present the phase synchrony analysis based on the Hilbert transform.

A. SSVEP enhancement using spatial filtering

SSVEP model: A signal recorded at a particular electrode location, that contains T samples can be seen as a vector in the space \( R^T \). Following this interpretation, we use hereafter the terms vector and signal without explicit distinction.
The signal \( x_i \), where \( i \) indexes the electrode location, recorded while the subject focuses his/her attention on a RVS modulated at a certain frequency \( f \), can be written as the sum of the SSVEP component (denoted as \( s_i \)), background EEG and noise [11]. For convenience, the background EEG and the noise at electrode \( f \) are combined into a single term denoted by \( y_i \). Thus, the following relation holds

\[
x_i = s_i + y_i,
\]

where the SSVEP component is modeled as a linear combination of vectors in the set

\[
\Phi = \left\{ \sin(2\pi ft), \cos(2\pi ft) \mid h = 1, \ldots, H \right\}.
\]

Here \( t = \left[ 0, T - 1 \right] \) is a vector of sample indices, \( H \) is the number of harmonics that are considered in the model, and \( a_{h,i} \) and \( b_{h,i} \) are real numbers. Equation 1 can be generalized to the whole set of electrodes \( \{x_i \mid i = 1, \ldots, N\} \) (\( N \) is the number of electrodes) in the following matrix form

\[
X = SA + Y,
\]

where the matrix \( X \in \mathbb{R}^{T \times N} \) (the notation \( \mathbb{R}^{T \times N} \) refers to the space of real matrices having \( T \) rows and \( N \) columns) has as columns the vectors \( x_i \), \( Y \in \mathbb{R}^{T \times N} \) has as columns the vectors \( y_i \), \( S \in \mathbb{R}^{T \times 2H} \) has as columns the vectors in the set \( \Phi \), and \( A \in \mathbb{R}^{2H \times N} \) is the matrix of linear combination coefficients such that: \( A_{h,i} = a_{h,i} \) is odd if \( h \) is odd and \( A_{h,i} = b_{h,i} \) if \( h \) is even. By means of the coefficients \( a_{h,i} \) and \( b_{h,i} \) the model in Equation 3 takes into account the differences of SSVEP-strengths across the scalp.

**Construction of the spatial filter:** The elements of \( A \) in Equation 3 cannot be determined from \( X \) and \( S \) only. Thus, to determine the SSVEP strength at different electrode locations, a signal \( x_w \) is constructed such that:

\[
x_w = \sum w_i x_i = Xw,
\]

where \( w = [w_1, \ldots, w_N]^\top \). The signal \( x_w \) can be considered to be the result of a spatial filter (e.g., filtering across the various electrodes) with coefficients \( \{w_i\} \) applied to the measured signals \( x_i \). The spatial filter is determined in such a way that it simultaneously maximizes the energy in the SSVEP frequencies and minimizes the energy in the nuisance signals [11]. This ratio can be determined by relying on the geometric interpretation described as follows.

The linearly independent vectors in the set \( \Phi \) generate a vector-space (\( \Pi \)) in \( \mathbb{R}^2 \) of dimension \( 2H \). In this article’s framework, we assume \( \Pi \) to be the space where the SSVEP components lie. Thus, \( \Pi \)'s orthogonal complement \( \Pi^\perp \) contains the non-SSVEP components.

Since the vectors in \( \Phi \) are linearly independent, the projection matrix \( Q \) on \( \Pi \) can be written as \( Q = S(S'S)^{-1}S' \) [12]. The component of \( Xw \) in \( \Pi^\perp \) is equal to \( Xw - QXw \). The Euclidean norm of the latter:

\[
\|Xw - QXw\|_2^2 = \text{divided by } T \text{ represents the power of the non-SSVEP related activity.}
\]

The power of the SSVEP related activity in \( Xw \) can be approximated by \( w'X'Xw \). The spatial filter \( w \) corresponds to the argument that maximizes the ratio

\[
\rho = \arg \max_{w} \frac{w'X'Xw}{\|X - QXw\|^2}.
\]

The ratio in Equation 4 is a generalized Rayleigh quotient [13] whose maximum can be found through a generalized eigen decomposition of the matrices \( X'X \) and \( (X - QX)'(X - QX) \). This results into two matrices \( W \), \( \Lambda \in \mathbb{R}^{N \times N} \) such that

\[
X'X = (X - QX)'(X - QX)WA,
\]

where \( \Lambda \) is a diagonal matrix whose diagonal contains the eigenvalues. The corresponding eigenvectors are in the columns of \( W \). By construction, eigenvalues are larger than one [14]. The largest element in \( \Lambda \) corresponds to the maximum of the quotient in Equation 4. The column of \( W \) corresponding to such maximum is the sought spatial filter \( w \).

In this study, we used the training strategy detailed in [10] to obtain the coefficients \( w \). Using these coefficients we obtained the single channel signal \( x(t) \)

\[
x(t) = Xw.
\]

**B. Phase synchrony analysis**

The phase difference between the SSVEP and the stimulus signal was extracted from the signal \( x(t) \), which is the bipolar combination of \( Oz \) and \( Cz \) (\( Oz-Cz \)) or the signal obtained from spatial filtering, and the light signal recorded by a photodiode. Both signals were recorded with a sampling rate of 2048 Hz and downsampled to 256 Hz. An FIR bandpass filter centered at the stimulus frequency \( f \) and having a bandwidth of 1 Hz was used to filter \( x(t) \) and the light signal to obtain the signals \( s(t) \) and \( l(t) \) respectively. The FIR filter was adopted because it can keep the phase linear. After having carried out the above procedure, we considered the obtained signal \( s(t) \) an approximation of the SSVEP.

The next step was to obtain the spectro-temporal representation of the two signals \( s(t) \) and \( l(t) \) such that the
instantaneous phase can be derived. This can be done using the Hilbert transform. The Hilbert transform takes a function and produces another function within the same domain so that an analytic representation (i.e. the spectro-temporal representation at the stimulus frequency $f^*$) of $s(t)$ or $l(t)$ is obtained. The Hilbert transform can be seen as the convolution of a signal with the function $h(t) = 1/\pi t$, such that

$$s_h(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} s(\tau) h(t-\tau) d\tau,$$

(7)

and

$$l_h(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} l(\tau) h(t-\tau) d\tau.$$

(8)

This means that the Hilbert transform can be realized by an ideal filter whose amplitude is unity and phase response is $\pi/2$ lag at all frequencies [15].

The analytic representations $S(f,t)$ and $L(f,t)$ of $s(t)$ and $l(t)$ are defined as follows

$$S(f,t) = s(t) + j \cdot s_h(t) = A_t(t) e^{j\phi_t(t)},$$

(9)

$$L(f,t) = l(t) + j \cdot l_h(t) = A_t(t) e^{j\phi_t(t)}.$$  

(10)

The phase difference $\Delta \phi_t(f,t)$ at the frequency $f$ and time $t$ between light and SSVEP can be obtained from $s(t)$ and $l(t)$

$$\Delta \phi_t(f,t) = \arctan\left(\frac{L(f,t)S(f,t)}{|L(f,t)||S(f,t)|}\right).$$

(11)

In this study, the mode of the distribution of $\Delta \phi_t(f,t)$ during a time window $[t_1,t_2]$ will be used as the desired phase difference $\Delta \phi$.

III. EXPERIMENTAL SETUP

In this study, two simultaneously oscillating green LED boxes (a power LED was mounted in a 10 cm × 10 cm box and the box was covered by a diffusion panel) driven by square wave currents rendered from two Agilent Function Generators (Model: 33220A) were used to present the stimuli. A luminance meter was used to measure the maximum luminance of two LEDs and the background luminance. In this experiment, the maximum luminance was 1714 nits and the background luminance was 69.68 nits. Therefore, the modulation depth of these two LEDs was (1714-69.68)/(1714+69.68) × 100% = 92.19%.

The distance between two LED boxes was around 30 cm and the subjects were seated on a chair located 70 cm from the LED boxes. These two LEDs oscillated at the same frequency $f$ but with different phases ( $\phi_1 = 0$, $\phi_2 = 2\pi/3$).

The optimal frequency $f$ for each subject was determined by a procedure of frequency selection. In this procedure, we presented four stimuli for each frequency from 40 Hz down to 31 Hz and simultaneously recorded the EEG signals. We calculated the energy at the frequency $f$ from the signal after spatial filtering (See section II-A) and used the receiver operating characteristic (ROC) curve to select the best frequency for each subject. With the ROC curve, we obtained the area under curve (AUC) for each frequency. The frequency was the best if the AUC was the largest. The selected frequencies for subjects are either 31 or 32 Hz, as shown in Table I. A photodiode was used to record the light source with $\phi_1 = 0$. The signal recorded from the photodiode was used in extracting the phase difference between the light signal and SSVEP using phase synchrony analysis (see section II-B).

Three male and one female subject participated in this study. Two subjects had normal vision, and the other two subjects had corrected to normal vision. They signed an informed consent before engaging in this study and had the right to quit at any time. Subjects were requested to pay attention to one of the two LEDs for 3 seconds defining 1 trial. Forty trials were performed consecutively with a rest period of a randomized duration between 3 and 5 seconds. In each trial, subjects were asked in turn to focus on one of the two LEDs, leading to 20 trials for each LED. The process was performed twice for each subject. Thus, each subject was presented with 40 stimuli for each LED.

EEG signals were collected using a BioSemi Active-two EEG acquisition device in a normal office environment with curtain closed and lights on. Thirty-two electrodes were placed according to the international 10-20 system. The impedance between scalp and electrodes was kept below 5 kΩ. The sample frequency was 2048 Hz. During the recording of EEG signals, subjects were advised to avoid movement or blinking during stimulus periods while being advised to blink during the rest period between two consecutive stimuli. They could move and rest between sessions.

In addition, to validate the fact that only a limited number of frequencies above 30 Hz can be used for BCI purposes, subject 1 and subject 2 participated in one more experiment in which the SSVEP energy of the frequencies from 30 Hz to 40 Hz would be observed. During this experiment, a subject would be presented with 10 trials for each frequency, which had the same timing (a stimulus period of 3 seconds and a rest period of a randomized duration between 3 and 5 seconds) as the previous experiment, followed by a rest period of at least 30 seconds when the frequencies changed. To reduce the subject’s fatigue, the frequency started from 40–Hz and were decreased to 30 Hz with 1 Hz steps.

IV. RESULTS

Figure 1 shows the phase difference between the SSVEP and the light signal extracted from 1-second EEG signals recorded from subject 1 while the subject focused his attention on one of the two LEDs using the phase synchrony analysis described in Section II. From the figure, we can see that the phase difference can be extracted by the phase
synchrony analysis as it fluctuates around a certain constant value. The measured phase difference values can, therefore, be used to identify which LED the subject was focusing on.

In our study, we examined the effect of spatial filtering on the phase of the SSVEP signal. We used the algorithm discussed in Section II to obtain the spatial filters for each subject. Figure 2 shows the topographical variability of the spatial filters for all subjects. The topography for each subject is different. This difference means that the spatial filter is subject-dependent.

Figure 3 shows the distributions of the phase differences of the EEG signal Oz-Cz (Figure 3(a)) and the signal after spatial filtering (Figure 3(b)) of subject 1. The electrodes used for spatial filtering were O1, O2, Oz, PO3, PO4, P3, P4 and Pz. Recorded EEG signals from 1 second after stimuli onset (i.e. the first second was ignored) of each trial were compared for each subject over each of the 80 trials. A 1-second window was used to calculate the features and there was no overlap between windows, i.e., we obtained 2 values for each trial. Figure 3 clearly indicates that the separability of two classes becomes better after spatial filtering.

In order to obtain the classification accuracy for all subjects, a support vector machine (SVM) [16] was used because it can adaptively determine the classification boundary and performs well for two classes. The first half of obtained phase difference values was used for training and the other half for testing. Table I shows the classification accuracy on the recorded data sets of four subjects. The results demonstrate that spatial filtering can increase the classification performance of the phase difference by enhancing SSVEP energy.

Figure 4 shows the SSVEP response to the stimuli ranging from 30 Hz to 40 Hz of subject 1 (Figure 4(a)) and subject 2 (Figure 4(b)). Ten energy values for each frequency were used to obtain the boxplot. The energy was calculated with a peak filter centered at the stimulus frequency and having a 1-Hz bandwidth. A 1-second window was used to obtain the averaged energy and there was only 1-point overlap between windows. The result shows that not all frequencies have a good SSVEP response. That is, the number of frequencies which can be used for SSVEP-based BCIs is limited.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Frequency</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Oz-Cz</td>
<td>Spatial Filtering</td>
</tr>
<tr>
<td>1</td>
<td>32 Hz</td>
<td>81.25%</td>
</tr>
<tr>
<td>2</td>
<td>32 Hz</td>
<td>98.75%</td>
</tr>
<tr>
<td>3</td>
<td>31 Hz</td>
<td>56.41%</td>
</tr>
<tr>
<td>4</td>
<td>32 Hz</td>
<td>66.25%</td>
</tr>
</tbody>
</table>

![Fig.1. Instantaneous phase difference extracted from a 1-second EEG segment having a good SSVEP response when subject 1 focused his attention on one of the two LEDs oscillating at 32 Hz. The EEG signal was from the bipolar combination of Oz and Cz.](Image)

![Fig.2. Topographical variability of the spatial filter w of the four subjects (the norm of w was 1).](Image)

![Fig.3. Distribution of the phase difference of subject 1: (a) calculated from Oz-Cz and (b) calculated from the one channel signal after spatial filtering.](Image)

![Fig.4. SSVEP energy of subject 1 (a) and subject 2 (b), these values were calculated from the signal Oz-Cz and they represent the median of the SSVEP energy of each trial. In each box, the central line is the median, the edges of the box are the 25th and 75th percentiles, and the whiskers extend to 10th and 90th percentiles, outliers (+) are plotted individually.](Image)

V. DISCUSSION

The phase synchrony analysis can effectively extract the phase difference between the SSVEP and the light signal, as shown in Figure 1. The difference between these two values deviates slightly from the expected value $2\pi/3$, but the separation is sufficient for classification. This deviation is probably caused by the strength of SSVEP response, the finite EEG points for calculation and the estimation error of the phase synchrony analysis.
In this study, the Hilbert transform was used to obtain the spectro-temporal representation of a signal. According to its definition, the Hilbert transform can be applied to any arbitrary signal to extract the instantaneous phase according to the definition. Nevertheless, the phase has a clear physical meaning only if the signal is a narrow-band signal. This is the reason why we use an FIR filter centered at the frequency of interest. Apart from the Hilbert transform, the Wavelet convolution using a Gabor function as the wavelet to obtain the spectro-temporal representation is also used for the analysis of neuronal synchrony [15]. The difference between these two methods is minor and they are fundamentally equivalent for the study of neuroelectric signals. The Hilbert transform, however, is less computationally intensive compared to the Wavelet convolution.

As shown in Table 1, spatial filtering can significantly increase the classification accuracy. This indicates that spatial filtering is important not only for estimating SSVEP energy [10] but also SSVEP phase. In this study, the coefficients of the spatial filter for each subject are fixed. Thus, the phase difference after spatial filtering is non-linearly and invariably related to the SSVEP phase of the electrode signals which are used to design a spatial filter.

The stimulus phase has the possibility to deviate from the preset value due to the design of stimuli if a long term presentation is used. Using the phase difference between the SSVEP and the stimulus signal as a feature can mitigate the adverse effect brought by this deviation, because SSVEP is phase-locked and the phase deviation of the stimulus causes the same phase deviation of SSVEP. Thus the subtraction of the phases of the SSVEP and the stimulus signal will cancel this deviation out. Here the stimuli rendered by both LEDs had the same change over time, because we used the same model of LEDs and function generators.

The SSVEP response of subject 1 and subject 2 in Figure 4 is different. This difference means that the SSVEP response is really subject-dependent. Furthermore, the classification accuracy of four subjects varies in a large range, as shown in Table I. This variation further indicates this dependency. Therefore, incorporating phase response in a frequency selection procedure should be considered in order to obtain better performance in future research.

VI. CONCLUSION

In this study, we used phase synchrony analysis to extract the phase difference between SSVEP and stimuli at frequencies above 30 Hz. We compared the classification performance using a single lead EEG signal (Oz-Cz) versus a one-dimensional signal from spatial filtering. Our result shows that phase synchrony can effectively extract the phase difference and spatial filtering can significantly increase the classification accuracy. Using the SSVEP phase is beneficial for increasing the available number of visual stimuli at higher frequencies (>30 Hz). Furthermore, the application of spatial filtering can increase the accuracy of determining a subject’s intention. Our upcoming work will focus on examining the usability of more than two phases for one frequency and testing the possibility of real-time applications in SSVEP-based BCIs.

REFERENCES