High Rate BCI with Portable Devices based on EEG

Zhengrui Qin[a,*], Qun Li[b]

[a] *School of Computer Science and Information Systems, Northwest Missouri State University, United States*
[b] *Computer Science Department, College of William and Mary, United States*

Abstract

The steady-state visual evoked potential (SSVEP) signal is widely utilized for Brain-Computer Interfaces (BCIs) that enable a direct communication between a user and an external device. Most devices in high-bit-rate SSVEP-based BCI, however, are lab-oriented and expensive, and thus not suitable for personal use in daily life. In this paper, we aim to investigate the feasibility of implementing high-bit-rate BCI using consumer-grade portable devices. We utilize a portable EEG headset to capture brain activities and extract user's intention from the weak and noisy signal through an optimization framework. We investigate several important factors that affect the bit rate and finally maximize the bit rate with the best choices of these factors. We can achieve 54.8 bits/minute on a normal computer monitor, which is comparable to the result with expensive devices, and 17.3 bits/minute on a smartphone screen, which is the first work of such on a smartphone screen. Our work can be easily integrated into security application such as defending against shoulder-surfing attacks, and assist patients with reduced motor abilities as well.

Keywords: EEG, SSVEP, Brain-Computer Interface, bite rate, canonical correlation analysis.

1. Introduction

There is currently much ongoing research in mobile health, body-area sensor networks, and wearable computing to sense human body and behaviors. As one of the most important organs, the human brain has drawn special attention. While there is an established category for Brain-Computer Interfaces (BCI), the full capabilities of these systems are yet to be realized, especially using portable devices only. Our vision of the future includes brain sensing systems that have a profound impact on daily life. Users will be able to read each others thoughts and emotions directly as a completely new method of communication. However, the first step to building such systems lies in our ability to extract information from the brain using current portable technology. This step allows us to begin understanding how the brain functions, and what is required to interface with it. In this paper we strive to build a system that senses the human brain using portable BCI devices. Our goal is to maximize the bit-rate between the brain and the physical world.

A BCI is a direct communication between the brain and an external device. For instance, a person who cannot physically use a keyboard may "type" by simply thinking of typing or gazing at a computer monitor. For example, the great physicist Stephen Hawking is cooperating with some scientists in NeuroVigil Inc. to establish faster communication through BCI technology [1]. Furthermore, BCI applications are recently emerging for a wide range of areas, such as gaming, entertainment, and secure channel. The information source used in BCIs is electroencephalography (EEG), which records the electrical activities of the brain through the use of sensors along the scalp. Several EEG signals, such as event-related synchronization/desynchronization, slow cortical potential, P300, visual evoked potential, and steady-state visual evoked potential (SSVEP), are commonly harvested for BCI control. Among these, SSVEP has received much attention due to several well-known advantages, such as high information transfer rate, short training time and simple system configuration. SSVEP is a brain response from the visual cortex in the rear scalp, induced by a repetitive visual stimulus that flickers at a constant frequency. In an SSEVP-based BCI, stimuli

*Corresponding author
Email addresses: zqin@nwmissouri.edu (Zhengrui Qin), liqun@cs.wm.edu (Qun Li)

with different frequencies are simultaneously displayed on a monitor, with each stimulus associated with a distinct symbol. To choose a symbol, a user gazes at the corresponding stimulus for a certain time, during which EEG signals are collected. By processing the collected EEG signal, the stimulus that the user is gazing at can be determined, so is the symbol the user intends to choose.

Current SSVEP-based BCIs, however, are built with specially designed devices. These devices are non-portable due to their huge size and unwieldy wired sensors, and mainly utilized in laboratories. Furthermore, these devices are very expensive; for example, Biosemi ActiveTwo devices [2] range from $20,000 to $100,000. Therefore, they are not suitable for personal daily use. In this paper, we aim to build a high rate SSVEP-based BCI system with portable devices, which has been seldom investigated so far. Our BCI system consists of an EEG device and a laptop or a smartphone, which are portable and friendly for personal use. The EEG device is an Emotiv headset [3], which costs only about $200. We believe headsets like this will become even cheaper if they are manufactured in mass.

When using portable devices, we find that the experience of portable devices is much different from that of expensive devices in many aspects, and we encounter many challenges and difficulties. As a result, we cannot apply existing work directly on BCI systems using portable devices only. First, compared to expensive devices, the cheap devices have much less sensors and sample at much lower rate. Second, due to the quality of the sensors and the loose contact, the EEG signal is much noisier and weaker than that from expensive devices. Third, during our experiments, we have found that some existing principles with expensive devices are not valid for cheap devices, which substantially complicates the entire system design. Finally, there are many factors that affect the bit rate, and it is non-trivial to find the optimal configuration of these factors to boost the bit rate. In summary, it is very challenging to build a high bit rate BCI system with portable devices.

We have managed to overcome these challenges and built a prototype of SSVEP-based high bit rate BCI system. In summary, we make the following contributions:

- We have built a prototype of the SSVEP-based BCI system with portable devices and have demonstrated the feasibility of obtaining high bit rate with such system.
- We have designed a novel optimization framework to process the weak and noisy raw data, which boosts the bit rate notably.
- We have identified and solved several challenges that come along only with portable devices.
- We have investigated several factors that affect the bit rate, and figured out the best configuration through extensive experiments.
- We have extensively evaluated our BCI system. Our system can achieve a bit rate of 54.8 bits/minute on a computer monitor, which is very close to the best result from expensive devices and almost doubles that from similar devices, and 17.3 bits/minute on a smartphone screen, which is the first work of such.

The rest of the paper is organized as follows. We review the related work in Section 2, and the architecture of our BCI system in Section 3. Sections 4 presents the result by implementing an existing work directly, showing that direct re-implementation dose not work with cheap devices, and describes three main challenges. Section 5, 6, 7 solve these three challenges, respectively. We present the evaluation in Section 8 and conclude the paper in Section 9.

2. Related Work

Mobile health, body-area sensor networks, and wearable computing are popular nowadays to sense human body and behaviors [4, 5, 6, 7]. In the following, we will review the category for human brain, called Brain-Computer Interfaces (BCIs).

BCI has been studied for several decades on the main purpose of assisting disable people. Most of BCI systems utilize the following EEG signals: visual evoked potentials (VEP), slow cortical potentials, P300 evoked potentials, sensorimotor rhythms, and cortical neurons. SSVEP, which is one of VEP, has received much attention due to several well-known advantages, such as high information transfer rate, short training time and simple system configuration. Researchers have built SSVEP-based BCI systems with expensive and non-portable devices with the goal of obtaining high information transfer rate (bits/minute). For instance, [8] reported a rate of 58 ± 9.6 bits/minute. [9] reported an

average bit rate of 61.70 ± 32.68 across seven subjects. Though the above works can achieve considerably high bit rates, they are non-portable and are not friendly for daily personal use.

For ordinary users, it is desirable to have portable devices, such as Neurosky [10] and Emotiv [3]. Only a handful of work has been done using cheap and portable devices. Campbell etc. [11] have designed and implemented a $NeuroPhone$ system using an Emotiv headset, in which a brain-controlled address book dialing app is demonstrated through a P300 BCI. Paper [12] has demonstrated the feasibility of side-channel attacks with an Emotiv headset. However, there is little work that uses these devices to implement SSVEP-based BCI, except [13, 14], where the bit rate is below 30 bits/minute.

In this paper, we aim to build a high-bit-rate SSVEP-base BCI system with portable devices on a laptop and a smartphone, respectively, with the goal of boosting the bit rate. When displaying stimuli on a laptop screen, our result nearly doubles that of [13] and [14]. For the scenario of using smartphone, we are the first to display stimuli on a smartphone screen to implement a BCI channel. Although in [13] the authors also utilized a smartphone, they just processed the EEG signal with a smartphone while the stimuli were still displayed on a large monitor. In our work, we use the smartphone's screen to display stimuli and process the EEG signal as well. Since the smartphone's screen is much smaller than a monitor, it is very challenging to obtain high bit rate.

It is worth noting that eye tracking has emerged as a promising technique to control an external device, one may think it could be an alternative. Compared with SSVEP based BCIs, the most significant barrier of eye track is to constrain the physical relationship between the participant and the eye tracking system. For instance, the researchers had to mount a restriction system on participant's head [15] to avoid tracking loss. Furthermore, eye tracking needs a bright environment and cannot be used in a dark room, while SSVEP-based BCIs do not have such constraints.

3. Architecture

Most of the existing BCI systems are built with expensive and non-portable devices, and are not ready for personal use in daily life. Our motivation is to bridge this gap by building a high bit rate BCI system with portable devices, and apply it to security scenario such as defending against shoulder-surfing attacks or use it to assist people with motor disability.

Our SSVEP-based BCI system is illustrated in Figure 1. The system has three physical components and three underlying components. The physical components are an EEG headset, a regular computer monitor or a smartphone, and an input application. The underlying components are *Data Acquisition*, *Signal Processing*, and *Command Translation*.

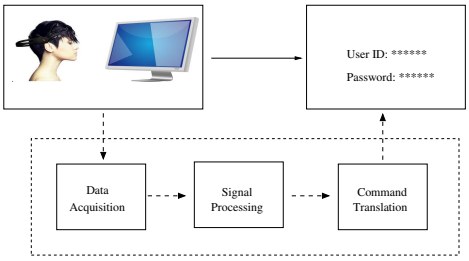

Figure 1: The architecture of SSVEP-based BCI.

The user runs an application that requires her to enter some characters such as user ID and password through her brain activities. She wears an EEG headset, facing a regular computer monitoror a smartphone displaying stimuli. To input a character, she gazes at the flickering stimulus associated with that character for a certain duration. EEG signal is collected by the Data Acquisition component in real time, then processed by the Signal Processing component, and finally translated to a character as the input by the Command Translation component.

3.1. Prototype System

We have built a prototype of our BCI system, as shown in Figure 2, which only consists of an Emotiv EEG headset and a Lenovo Windows 7 laptop or a smartphone. The EEG headset used in our BCI system is an Emotiv Epoc headset [3], as shown in Figure 3. There are 16 sensors in total, two of which are reference sensors. As SSVEP signals are generated from the rear region of the scalp, only four sensors, located at P_7, P_8, O_1, and O_2 in the international 10-20 system notation, are effective in our BCI system.

Compared to BCI systems with expensive and laboratory-oriented devices, our BCI system obviously has two advantages. One is portability; our system is convenient for personal use in daily life. The other is low cost; our system is only about 1/100 of the laboratory-oriented system in price.

4. Challenges

Since BCI systems with expensive and non-portable devices have been intensively studied, one would wonder whether the system of this kind can be directly applied to portable devices. To answer this question, we have implemented an existing system by directly substituting expensive devices with portable devices, through which we have identified several challenges particularly related to portable devices.

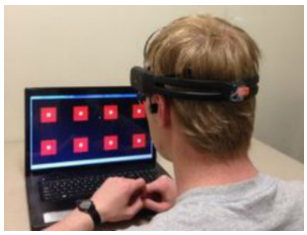

Figure 2: Our prototype system.

Figure 3: Emotiv headset.

4.1. Re-Implementation

In paper [8], researchers have built an SSVEP-based BCI using an LCD monitor (VIEWSONIC, 17", 60 Hz refresh rate) and a BioSemi ActiveTwo EEG headset (at least $20,000) sampling at a rate of 256 Hz. Six boxes are arranged in the form of 2×3 grid on the monitor. The six flickering frequencies are 15 Hz, 12 Hz, 10 Hz, 8.6 Hz, 7.5 Hz, and 6.7 Hz, which are divisible to the monitor refresh rate (60 Hz). The final bit rate[1] is 58 ± 9.6 bits/minute.

We have re-implemented exactly what was conducted in [8], except that we use the much cheaper and portable Emotiv headset. The final bit rate is 17.9 bits/minute. In fact, even this result may be inflated. We found that the 0.3 second interval between choosing two symbols in [8] was not long enough for the subjects to shift their focuses. During this interval, besides shifting their focuses, the subjects must also think about the next symbol they want to input. In our experiments, the subjects are instructed to gaze at the targets in an fixed order, and the subject actually do not need time to think about which symbol to focus on next. Based on what we have experienced, a one second interval of rest is comfortable for subjects to shift their focuses. Considering this, the bit rate would reduce to 13.7 bits/minute.

The main reason for the big decrease in bit rate between the non-portable and portable devices is the identification accuracy. We thus calculated and listed the accuracy for each individual target in our re-implementation, as shown in Table 1. We can see that the accuracy rate in the low frequency range is higher than that in the high frequency range. The highest rate, 96%, occurs at the target flickering at the rate of 7.5 Hz, and the lowest one, 0%, occurs at the target flickering at the rate of 15 Hz.

Table 1: Identification accuracy for each target (frequency).

Frequency (Hz)	15	12	10	8.6	7.5	6.7
p (%)	0	60	60	60	96	83

4.2. Challenges

Compared with expensive devices, we find that our SSVEP-based BCI system with portable devices faces three major challenges based on the re-implementation and the product characteristics of the Emotiv headset.

- The EEG signals from portable devices are weaker and noisier. The Emotiv EEG headset has only 4 effective sensors[2] and much lower sampling rate (128 Hz), while Biosemi ActiveTwo has 32 effective sensors and samples up to 1000Hz. Furthermore, the contact between Emotiv sensors and the scalp is much looser since Emotiv headset uses dry sensors while Biosemi applies lotion on the scalp.

- The effective frequency range is much narrower. As shown in Table 1, 15 Hz does not work at all. The narrow effective frequency range complicates the entire frequency assignment for SSVEP substantially.

[1] The bit rate is the actual information transfer rate where errors have already been corrected, and thereinafter.

[2] Even though Emotive headset has 16 sensors, most of them are not SSVEP-effective. Due to the fixed arms of the sensors, only 4 of them can be placed in SSEVP-effective area of the scalp.

- Multiple factors affect the bit rate, such as the number of stimuli, the size of stimuli, the color of stimuli, etc. It is challenging to configure these factors to achieve an optimal bit rate.

In the following sections, we will investigate all these challenges and present solutions for them.

5. Data Processing

In this section, we will design and describe our algorithm to process the weak and noisy EEG signal. In our algorithm, the raw EEG signals will first be pre-processed, and then be fed into an optimization framework that identifies the target the user is gazing at.

5.1. Signal Pre-processing

We notice that the contact between the sensors and the scalp is not perfectly tight due to the user's hair and dry sensors (with no gel). During signal collection, a small movement, such as head moving, eye blinking, even swallowing and breathing, may change the contact intensity, leading to dramatic amplitude change in the EEG signal. Our pre-processing algorithm is a high pass filter, which filters out the linear trend of the raw signal that results from the contact variation, which enables us to focus only on the fluctuations that reflect the repetitive flickering, i.e., the frequency information.

5.2. Optimization Framework

Before detailing the optimization framework, we will first explain a fundamental component of the framework, canonical correlation analysis (CCA). CCA is a multivariable statistical approach to analyze the underlying correlation between two sets of data. It has been used in EEG processing previously [16, 8, 14]. Given two sets of data, CCA determines two linear combinations for these two sets of data respectively such that the correlation between the two linear combinations is maximized. In our EEG signal processing, the first set of data is the EEG signal, denoted by $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4)$, where $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$, and $\mathbf{x}_4$ are time-series signals sampled at frequency F_s from senors O_1, O_2, P_1, and P_2, respectively. The other set of data is a series of sine and cosine waves with a fixed frequency, denoted by $\mathbf{Y} = (sin(2\pi f\mathbf{t}), cos(2\pi f\mathbf{t}), sin(4\pi f\mathbf{t}), cos(4\pi f\mathbf{t}), ..., sin(2h\pi f\mathbf{t}), cos(2h\pi f\mathbf{t}))$, where h is the number of harmonics, f is a fixed frequency, and $\mathbf{t}$ is the time vector at which $\mathbf{X}$ is sampled. Let $\mathbf{x}$ and $\mathbf{y}$ be the linear combination of $\mathbf{X}$ and of $\mathbf{Y}$, respectively, i.e., $\mathbf{x} = \mathbf{X}\mathbf{w}_x^T$ and $\mathbf{y} = \mathbf{Y}\mathbf{w}_y^T$, where $\mathbf{w}_x = (w_{x_1}, ..., w_{x_4})$ and $\mathbf{w}_y = (w_{y_1}, ..., w_{y_{2h}})$. CCA aims to maximize the canonical correlation ρ between $\mathbf{x}$ and $\mathbf{y}$:

$$\begin{aligned} max_{\{\mathbf{w}_x, \mathbf{w}_y\}}: \quad \rho(\mathbf{x}, \mathbf{y}) &= \frac{\mathbf{x}^T\mathbf{y}}{\sqrt{(\mathbf{x}^T\mathbf{x})(\mathbf{y}^T\mathbf{y})}} \\ &= \frac{\mathbf{w}_x\mathbf{X}^T\mathbf{Y}\mathbf{w}_y^T}{\sqrt{(\mathbf{w}_x\mathbf{X}^T\mathbf{X}\mathbf{w}_x^T)(\mathbf{w}_y\mathbf{Y}^T\mathbf{Y}\mathbf{w}_y^T)}} \end{aligned} \tag{1}$$

The larger the value of ρ, the more likely $\mathbf{x}$ and $\mathbf{y}$ have the same frequency. Suppose there are N stimulus frequencies, denoted as $f_i, \forall i \in [1, N]$. For each stimulus frequency f_i, we can get a $\mathbf{Y}_i$, $\forall i \in [1, N]$. Given an $\mathbf{X}$ and a $\mathbf{Y}_i$, we can get a ρ_i using the CCA algorithm. Finally, we can conclude that the user is very likely gazing at the box flickering with frequency f_{i^*}, where $i^* = \arg\max_{i \in [1,N]} \rho_i$.

Based on the re-implementation of the work in [8], we find that the CCA result from a single chunk of EEG data performs poorly. To improve identification accuracy, we use a t_0-$\triangle t$ sliding-window method. We run CCA algorithm for each t_0-long signal, with $\triangle t$-long sliding window, until some pre-defined criteria are reached.

The objective function in our optimization framework is to maximize the information transfer rate [17] on average, i.e.,

$$max: \quad \frac{60}{\overline{s}}[log_2 N + plog_2 p + (1-p)log_2\frac{1-p}{N-1}] \tag{2}$$

where N is the number of flickering boxes on the monitor, $\overline{s}$ is the average time duration for one session (gazing at one box), and $p = n_0/n$ is the identification accuracy rate where n is the total number of sessions and n_0 is the number of sessions that are correctly identified.

An identification decision is made (at the end of each session) when the identified target dominates in multiple windows:

$$Max(\sum_{j=1}^{i-1} \boldsymbol{\rho}_j^{\alpha}) < \beta \times Max_2(\sum_{j=1}^{i-1} \boldsymbol{\rho}_j^{\alpha}) \tag{3}$$

$$Max(\sum_{j=1}^{i} \boldsymbol{\rho}_j^{\alpha}) \geq \beta \times Max_2(\sum_{j=1}^{i-1} \boldsymbol{\rho}_j^{\alpha}) \tag{4}$$

where $\alpha > 1$ and $\beta > 1$ are two variables, Max_2 is a function that returns the second largest element in a vector, i is the window index at end of which an identification decision is made (therefore, the session time $s = t_0 + (i-1)\triangle t$)), and $\boldsymbol{\rho}_j$ is the normalized vector obtained from window j through CCA algorithm, i.e.,

$$\boldsymbol{\rho}_j = [\rho_{j1}, \rho_{j2}, ..., \rho_{jN}] \tag{5}$$

$$\sum_{j=1}^{i} \boldsymbol{\rho}_j^{\alpha} = [\sum_{j=1}^{i} \rho_{j1}^{\alpha}, \sum_{j=1}^{i} \rho_{j2}^{\alpha}, ..., \sum_{j=1}^{i} \rho_{jN}^{\alpha}] \tag{6}$$

The variable α is equivalent to assigning non-linear weights to each element of vector $\boldsymbol{\rho}_j$. Since $\alpha > 1$ and $\rho_{jk} \leq 1$, $k \in [1, N]$, the greater the value of ρ_{jk}, the more the weight assigned to f_k. Eq (3) and Eq (4) indicate that an identification decision is made once the largest of the accumulated weighted ρ values is greater than β times of the second largest for the first time. We use brute-force search to find the best combination of α and β.

Algorithm 1: Frequency Identification Algorithm.

Input: h, t_0, $\triangle t$, f_1, f_2, ..., f_k, α, β, F_s;
Output: f_{out}, the identified frequency.

1 $i = 0$, $\boldsymbol{\rho}_{sum} = \mathbf{0}$, $\rho_{max} = \rho_{max2} = 99$, //initialization;
2 $\mathbf{t} = (\frac{1}{Fs}, \frac{2}{Fs}, ..., t_0)^T$;
3 Wait for t_0 after boxes start flickering at $t = 0$;
4 **while** $\rho_{max} < \beta * \rho_{max2}$ **do**
5 $\quad \mathbf{X}$ = 4-sensor signals from $i\triangle t$ to $i\triangle t + t_0$;
6 $\quad$ **for** $j = 1, k$ **do**
7 $\quad\quad \mathbf{Y} = (sin(2\pi f_j \mathbf{t}), cos(2\pi f_j \mathbf{t}), sin(4\pi f_j \mathbf{t}), cos(4\pi f_j \mathbf{t}), ..., sin(2h\pi f_j \mathbf{t}), cos(2h\pi f_j \mathbf{t}))$;
8 $\quad\quad \rho_j = max_{\{\mathbf{w}_x, \mathbf{w}_y\}} \frac{\mathbf{w}_x \mathbf{X}^T \mathbf{Y} \mathbf{w}_y^T}{\sqrt{(\mathbf{w}_x \mathbf{X}^T \mathbf{X} \mathbf{w}_x^T)(\mathbf{w}_y \mathbf{Y}^T \mathbf{Y} \mathbf{w}_y^T)}}$;
9 $\quad$ **end**
10 $\quad \boldsymbol{\rho} = [\rho_1, \rho_2, ..., \rho_k]$;
11 $\quad \boldsymbol{\rho} = \boldsymbol{\rho} / max(\boldsymbol{\rho})$, // normalizing;
12 $\quad \boldsymbol{\rho} = \boldsymbol{\rho}^{\alpha}$, // assigning non-linear weights;
13 $\quad \boldsymbol{\rho}_{sum} = \boldsymbol{\rho}_{sum} + \boldsymbol{\rho}$;
14 $\quad \rho_{max} = max(\boldsymbol{\rho}_{sum})$;
15 $\quad \rho_{max2} = max(\boldsymbol{\rho}_{sum} \setminus \rho_{max})$, // second largest;
16 $\quad i^* = \arg\max_{j \in [1,k]} \boldsymbol{\rho}_{sum}$;
17 $\quad i = i + 1$;
18 **end**
19 $f_{out} = f_{i^*}$;
20 Stop flickering for all boxes and prepare for next session;

Algorithm 1 details the entire optimization framework. lines 5-9 carry out CCA algorithm for each candidate frequency f_j and obtain the corresponding correlation ρ_j, $\forall j \in [1, k]$. Line 11-13 sums up the weighted normalized ρ values for all windows so far. Line 14 and Line 15 obtain the most dominant candidate and the second dominant candidate, respectively. The while-loop (line 4 to line 18) terminates when the most dominant candidate surpasses the second dominant candidate by a predefined ratio. After the target frequency is identified, all boxes stop flickering and prepare for the next session.

6. Frequency Assignment

Traditional wisdom, such as [8], tells us that the stimulus frequency for SSVEP must be divisible to the monitor's refresh rate and less than 20 Hz. For instance, for a monitor with a 60Hz refresh rate, the possible candidates could be 6 Hz, 6.7 Hz, 7.5 Hz, 8.6 Hz, 10Hz, 12Hz, and 15 Hz. From Table 1, however, we see that high frequency candidates perform poorly for portable devices. In this section, we will investigate the feasibility of using more frequency candidates and design proper frequency assignment for stimuli on the monitor.

6.1. Divisibility is not necessary

Through extensive experiments we find that the requirement of divisibility is not necessary. In the experiments, we use a Lenovo laptop running Windows 7, with a $17''$ monitor. There are two choices for monitor refresh rate, 40Hz and 60Hz, among which we choose the latter. We directly use a JAVA timer to control the stimulus frequencies. For instance, to obtain a frequency of 7 Hz, we let the stimulus alternate with two colors, each with a duration of $\frac{1}{14}$ seconds.

Surprisingly, we find that many frequencies can be identified through a simple FFT, even though they are not dividable to the refresh rate. We have tested frequencies ranging from 5Hz to 15Hz with 0.2Hz interval. For the spectrum from 5.6 Hz to 9 Hz, we can clearly see a peak at the corresponding frequency in the FFT result. Figure 4 illustrates one example, in which the top panel is the 8 second raw EEG data from a 7 Hz stimulus and the bottom panel is the FFT result after filtering out low frequency noise. The amplitude at 7 Hz is at least twice as large as any other peak. 7 Hz, however, is not dividable to the refresh rate 60 Hz.

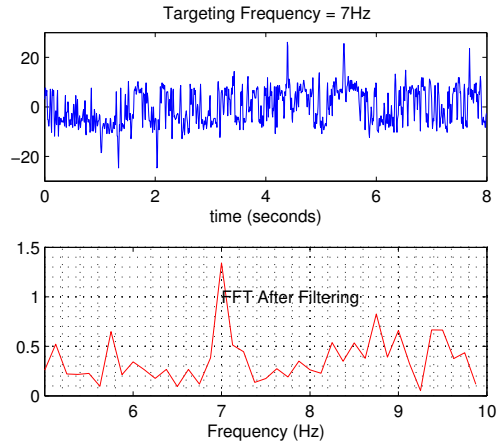

Figure 4: Original signals and its FFT for stimulus flickering at 7Hz.

6.2. Effective Frequency Range

From Table 1 we know that the effective frequency range will be narrower with portable devices, but it is still unclear what range is most effective. We will use the CCA algorithm to find the effective frequency range.

We ran intensive experiments with different flickering frequencies, ranging from 4 Hz to 15.2 Hz with a step size of 0.4 Hz. For each frequency, we run 180 rounds, in each of which the target flickers for two seconds. For each of the two second signals, we run CCA for each of the frequency candidates ($f = 4 : 0.4 : 15.2$) to identify the embedded frequency. If the identified frequency based on CCA and the embedded frequency of the target are same, we say the identification is correct; otherwise, it is wrong. The accuracy is shown in Figure 5. We can see that the accuracy is above 50% from roughly 5.6 Hz to 9.2 Hz. It is clear that there is a decreasing trend toward both higher frequencies and lower frequencies. We conclude that for portable devices like Emotiv headsets, the effective frequency range is much narrower than that for laboratory-oriented devices.

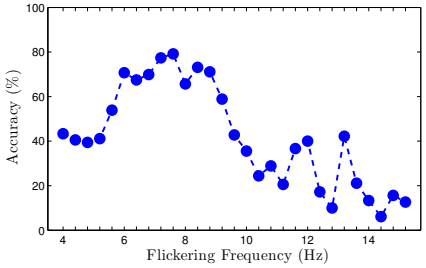

Figure 5: The identification accuracy with different flickering frequencies.

6.3. Frequency Gap

We also investigate the impact of the frequency difference or gap among the frequency candidates on the accuracy. Intuitively, according to the essence of CCA, the wider the gap, the higher the accuracy. The reason is that, a wider gap can result in a larger difference in the ρ values among candidate frequencies. However, since the effective frequency range is bounded (around 5.6 Hz to 9.2 Hz), we cannot choose a very big gap. As Figure 6 shows, the accuracy increases and then decreases when the gap goes from 0.1 Hz to 0.6 Hz. There is no need go beyond 0.6 Hz due to the limited effective frequency range.

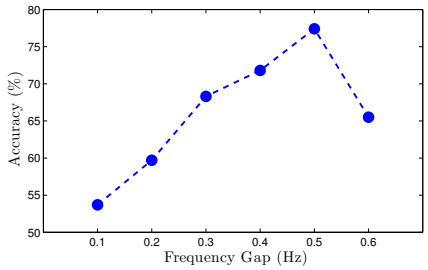

Figure 6: The identification accuracy with different frequency gaps.

6.4. Interference

Here we aim to determine whether the identification accuracy of one stimulus is affected by its neighboring stimuli, i.e., whether there is interference among the flickering targets. If the interference does exist, we must learn the intensity and the nature of the interference. In the experiment, we display 9 flickering boxes on the monitor in a form of 3×3 grid, as shown in Figure 7. The width of each box is 180 pixels, with a 40 pixel margin between boxes.

It is obvious that the distance between flickering targets will affect the identification accuracy. The larger the distance, the weaker the interference. Once the monitor is chosen and the number of targets is determined, we just need to separate them as far as possible. Therefore, here we only test the influence caused by neighboring targets' flickering frequencies. In this experiment, the subject is gazing at the center box, which is flickering with frequency f_0. The surrounding eight boxes are all flickering with the same frequency f_1. Let $\triangle f = f_1 - f_0$. We calculate the identification accuracy of f_0 by setting different $\triangle f$. As shown in Figure 6, the best frequency gap is 0.5Hz. Therefore, we are interested in the identification accuracy when $\triangle f$ is multiple of 0.5Hz. The results are shown in Table 2. As we can see, the identification accuracy is increasing when $\triangle f$ goes from 0.5Hz to 1.5Hz. That is, the accuracy has an increasing trend when $\triangle f$ increases. Therefore, once the frequency gap is determined, we need to assign flickering frequencies to boxes such that the frequency difference among neighboring boxes is as large as possible.

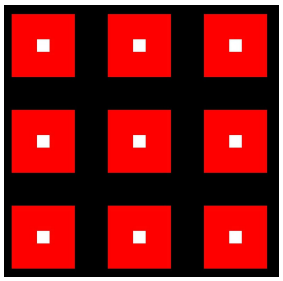

Figure 7: Interference setting.

6.5. Frequency Assignment

Based on the previous observations, our task here is to assign flickering frequencies to targets on the monitor. Suppose there are $m \times n$ targets placed on the monitor, where m is the number of rows and n is the number of columns. Now each target will be assigned a unique frequency chosen from a frequency set $\mathbf{f} = \{f_1, f_2, ..., f_{mn}\}$ with cardinality mn. Based on the observation on frequency gap, the effective frequency range $[f_{min}, f_{max}]$ is equally divided to obtain $\mathbf{f}$; that is,

Table 2: Identification accuracy under the interference of neighboring targets.

$\triangle f$	0.5	1.0	1.5
p (%)	59.3	63.9	69.0

$$f_i = f_{min} + \frac{(i-1)(f_{max} - f_{min})}{mn - 1}, \quad \forall i \in [1, mn] \tag{7}$$

Based on the observation of interference, the larger the frequency gap between a target and its neighboring targets, the less the interference. At the same time, the bigger the distance between two targets, the less the interference. Therefore, we assign a weight to each frequency gap, which is inversely proportional to the distance of two targets to whom the pair of frequencies are assigned. Our goal is to maximize the overall weighted frequency gaps among all targets. Since two targets far away from each other have a little or no interference, we only consider the direct

neighbors of each target. Each target has at most 8 direct neighbors. The frequency assignment problem is formulated as follows.

$$maximize: \quad \sum_{i=1}^{mn}(\sum_{j\in N_i}|f_i - f_j|/d_{ij}) \tag{8}$$

where N_i is the set of direct neighbors of box i, and d_{ij} is the distance between box i and box j.

In our implementation, $m = 2$ and $n = 4$ (to be explained later), $f_{min} = 5.6$, and $f_{max} = 9.1$. The spacing in rows is 1.2 times of that in columns. Since m and n are small, we resort to brute-force search to obtain the optimal frequency assignment. Table 3 shows one of the optimal assignments.

7. Configuration Design

In addition to the frequency assignment, there are several other factors that affect the bit rate, including the number of targets on the monitor, the size of the targets, the decision parameters in Algorithm 1, the distance between the user's eyes and the monitor, target pattern and color, etc. In this section, we assume a simplified model that these factors do not interact with each other. Under this simplified model, we can search the optimal or quasi-optimal configuration of each factor one by one by fixing all other factors.

Table 3: One optimal frequency assignment.

8.1	5.6	7.6	8.6
6.1	7.1	9.1	6.6

7.1. Impact of Number of Boxes

According to Eq (2), the number of targets, N, plays an important role in bit rate. One would think the more the targets, the higher the bit rate. However, as the number of targets increases, the interference also increases. At the same time, the frequency gap among targets decreases due to bounded effective frequency range. As a result, the accuracy rate will decrease. Here, we will calculate the bit rate with different number of boxes and determine the best number of boxes.

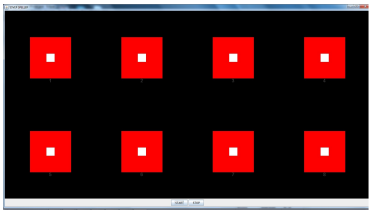

Figure 8: The layout of 8 boxes.

Using a 900×1500 pixel monitor, we did experiments with 2 boxes, 4 boxes, 6 boxes, 8 boxes, 12 boxes and 16 boxes. The layout of the boxes are 1×2, 2×2, 2×3, 2×4, 3×4, and 4×4, respectively, where $x \times y$ means that there are x rows and y columns of boxes. Figure 8 shows the layout of 8 boxes. On the screen, the background is black. Each box has an inner square and a square ring around the inner square. The outer ring and the inner square have different colors, either red or black, alternating with a fixed frequency. At the bottom, there is a Start button and a Stop button. For 8 boxes, the frequencies assigned is listed in Table 3.

We examine the bit rate (bits per minute), which is the ultimate metric of performance. Figure 9 shows the relation between the bit rate and the number of boxes. We can clearly see the increase-decrease trend. Only when a proper N is used, neither too small nor too large, does the optimization framework result in a high bit rate. Clearly, $N = 8$ is the best choice based on Figure 9. We do not consider 7 boxes, 9 boxes, 10 boxes, or 11 boxes, by considering the layout configuration and user experience.

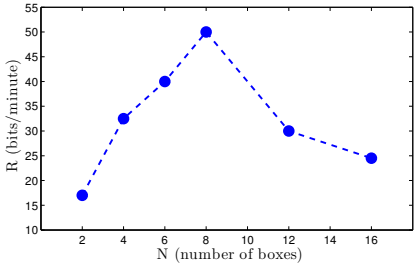

Figure 9: Bit rate (bits/minute) with respect to number of boxes.

7.2. Impact of Stimulus Size

The performance of SSVEP-based BCI mainly relies on the stimulus intensity. On one hand, if the boxes are too small, the stimulus will not be strong enough. On the other hand, if the boxes are too big, severe interference will be introduced into neighboring boxes.

We have conducted experiments to determine the best value of box size. We run several experiments many times with different values of box size, and evaluate the bit rate. In these experiments, we set the number of boxes $N = 8$, and the layout is shown in Figure 8. For data processing, we set 2 seconds as the window size and 0.25 seconds as the sliding size.

The bit rate against box size is shown in Figure 10. Based on this set of experiments, we set 180 pixels as the box size in our final implementation.

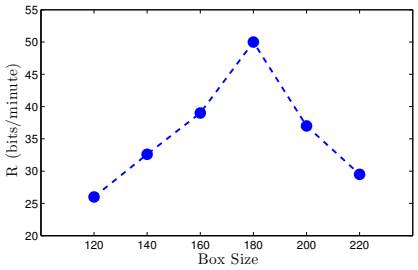

Figure 10: Bit rate (bits/minute) with respect to box size.

7.3. Impact of Window Size and Sliding Size

In previous experiments, we used a 2 second window size (t_0) and a 0.25 second sliding size ($\triangle t$) (see the definitions in Section 5.2). Here we explain why these parameters are chosen throughout our experiments.

We processed the data with different window sizes and sliding sizes. We set the window size ranging from 1 second to 3 seconds with a 0.25 second step, and set the sliding size of the sliding window ranging from 0.125 seconds to 1 second with a 0.125 second step. By doing this, we can obtain the bit rates (bits per minute) for different window sizes with different sliding sizes. Figure 11 show the bit rate. We can see that, if we fix the sliding size, the bit rate first increases and then decreases as the window size increases; if we fix the window size, the bit rate also first increases and then decreases as the sliding size increases. The bit rate peaks exactly at the position when the window size is 2 seconds and the sliding size is 0.25 seconds. Therefore, in our final implementation, we set 2 seconds as the window size and 0.25 seconds as the sliding size.

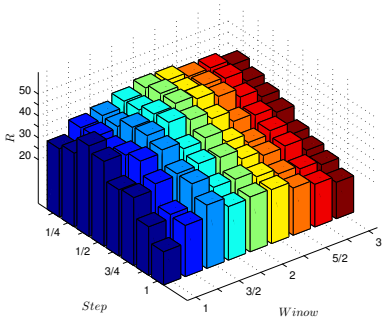

Figure 11: Bit rate (bits per minute) with respect to window size and sliding size.

7.4. Distance

The identification accuracy of SSVEP-based BCI depends heavily on the strength of the EEG signal, i.e., the quality that the retina responds to the flickering stimulus. It is obvious that the strength of the EEG signal is greatly influenced by the distance between the eyes and the monitor and by the angle between the line of sight and the plane of the monitor. As for the angle, it is clear that 90 degree is the best; therefore, when doing experiments, we always place the monitor such that the line between the center of two eyes and the center of the monitor is perpendicular to the monitor plane. Thus, this experiment solely focuses on the distance between the eyes and the monitor. We do not consider extreme cases, such as very small distance and very large distance. In this experiment, we measure the identification accuracy with distances ranging from 40 cm to 80 cm, with an interval of 10 cm.

The monitor layout is shown in Figure 8. There are eight boxes, flicking with frequencies of [5.6, 6.1, 6.6, 7.1, 7.6, 8.1, 8.6, 9.1] Hz, respectively. For each distance, the subject is instructed to focus on each box for 2 seconds for at least 90 rounds. We use the CCA algorithm to calculate the identification accuracy on average for each distance. The result is shown in Figure 12.

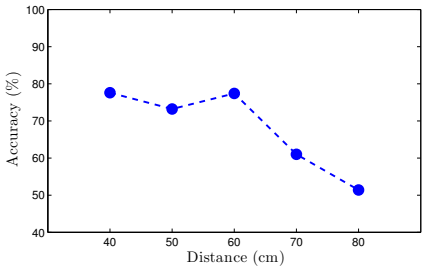

Figure 12: Accuracy vs Distance.

As we can see, from 40 cm to 60 cm, the identification accuracy is pretty similar, around 75%. When the distance keeps increasing, the identification accuracy drops sharply. Based on this observation, we decide to choose 60 cm as the distance in the final setting. There are two reasons. First, the identification accuracy is acceptable at 60 cm. Second, 60 cm is a normal distance between the eyes and the monitor when the user usually uses a computer.

7.5. Impact of Color and Box Pattern

We also conduct experiments for stimuli with different colors and with different patterns. The colors we have investigated are red, yellow, blue, black, and white. We find through experiments that the combination of red and white on a black background performs best. Besides the pattern shown in Figure 8, we have also examined the patterns shown in Figure 13. However, the performance of all patterns in Figure 13 are consistently worse. Our current box pattern has a small square box in the middle and a square ring outside of the inner square box. This pattern has at least two advantages. First, the inner small box can help users focus on the target; when users are doing experiments, they are instructed to gaze at the center of the target. Second, the inner part and the outer part alternate with red and white, which can intensify the stimulus.

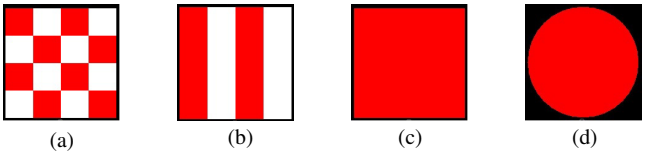

Figure 13: Stimulus patterns: (a) small squares; (b) vertical strips; (c) single square; (d) circle. The screen background is black. Each pattern alternates between two different colors, which are red and white in the current figure.

7.6. Configuration Summary

Based on extensive experiments, our final configurations are: the number of targets $N = 8$; target size = 180 pixels; distance between user's eye and the monitor = 60 cm; window size $t_0 = 2$ and sliding size $\triangle t = 0.25$ seconds. The colors for the targets are white and red. And the target pattern is a square box with a square ring outside.

8. Evaluation

In this section, we will present the evaluation of our prototype BCI system. First, we will examine the performance when stimuli are displayed on a computer monitor, including precision and recall, and online results. Second, we will show the performance when stimuli are displayed on a smartphone screen.

8.1. On Computer Monitor

8.1.1. Precision and Recall

Precision and recall are two important metrics for identification accuracy. The precision on one target is the ratio of the number of correct identifications of that target over the total number of sessions that are identified as that target including both correct and incorrect identifications. The recall on one target is the ratio of the number of correct identifications of that target over the total number of sessions the user has gazed at that target. Here we are interested in the precision and recall of our optimization framework. Table 4 shows the results of one subject. We can see that the precision and the recall across different frequency candidates do not fluctuate much. The average of the precision is 0.87, with a standard deviation of 0.04.

Table 4: Precision and Recall.

Frequency (Hz)	5.6	6.1	6.6	7.1	7.6	8.1	8.6	9.1
Precision	0.80	0.87	0.87	0.90	0.93	0.86	0.86	0.85
Recall	0.88	0.84	0.83	0.89	0.93	0.82	0.89	0.88

8.1.2. Online Results

We have recruited 4 subjects to conduct the online experiments. Before starting this work, we submitted a human subjects proposal to the Protection of Human Subjects Committee in the university and received a written approval. Before conducting experiments, we formally told each subject what the task was and explained there was no risk during experiments. Additionally, each subject signed a consent form. The age of the subjects ranges from 25 years to 36 years. All of them have short hair and (corrected) normal vision. The results are shown in Table 5.

Table 5: The online bit rate for four subjects.

	time (s)	p (%)	bits/min	bits/min*
subject 1	3.25	86.5	37.8	48.2
subject 2	3.09	95.0	50.0	64.6
subject 3	3.14	88.8	41.7	53.7
subject 4	3.04	86.5	40.4	52.5
average	3.16	87.7	42.5 ± 5.3	52.5 ± 7.0

The average bit rate is 42.5 bits/minute, with a standard deviation of 5.3 bits/minutes. To the best of our knowledge, the average bit rate with expensive devices is about 60 bits/minute, such as [8]. Note that in [8], the interval of rest between two sessions is 0.3 seconds, while it is 1 second in our BCI system. To make a fair comparison, we should calculate the bit rate by using the same interval. The last column in Table 5 shows the results after adjusting the interval time. The average bit rate becomes 54.8 bits/minute, with a standard deviation of 7.0 bits/minute. Then our result is very close to that in [8]. Recall the average bit rate from direct re-implementation in Section 4.1 (17.9 bits/minute), our result here is three times better.

Table 6: The improvements contributed by the pre-processing and the optimization framework.

	w/o both	w/o opt	w/o pre	with both
subject 1	38.9	41.1	47.0	48.2
subject 2	53.7	56.2	62.8	64.6
subject 3	41.7	49.7	47.9	53.7
subject 4	45.0	47.7	51.1	52.5

We also investigated the individual improvement contributed by each component of our algorithm. The base line is the algorithm in [8] that identifies the targets using the CCA algorithm with a 2 second data chunk. Our algorithm has two additional components, the pre-process and the optimization framework. Table 6 shows the improvements by these two components (the interval time of rest has been adjusted). "w/o both" is the base line algorithm. "w/o opt" is our algorithm without the optimization framework. "w/o pre" is our algorithm without the pre-processing. "with both" is exactly our algorithm in this paper. On average, the pre-processing improves 8.9%, the optimization framework improves 16.5%, and the combination of the two improves 22.4%.

8.2. On Smart Phone

Currently, many people use smartphones to conduct various online activities, which require user name and password. This may also introduce risk of shoulder-surfing attacks. Therefore, we also investigate the feasibility of SSVEP on smartphones. Since the screen of a smartphone is much smaller than that of a laptop, we cannot place as many targets as before. By repeating the same procedure as before, we finally display three squares on the screen of smartphone, as shown in Fig. 14. The frequencies used are 6 Hz, 10 Hz, and 8 Hz respectively.

We have conducted online evaluations with four subjects. The average bit rate is 17.3 bits/minute with 0.3 second rest interval, with a standard deviation of 2.7 bits/minutes, as listed in the last column of Table 7. Since there are only three targets, 0.3 second interval is enough to choose one from three. Anyway, we also show the results with 1 second interval in the fourth column.

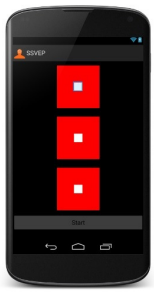

Figure 14: The configuration of SSVEP on a smartphone.

9. Conclusion

In this paper, we have carefully designed an SSVEP-based BCI system with cheap and portable devices, which can achieve high bit rate. The system can be integrated into security

Table 7: The online bit rate for four subjects on smartphone.

	time (s)	p (%)	bits/min	bits/min*
subject 1	3.31	85.2	15.1	19.1
subject 2	3.37	81.6	12.7	16.0
subject 3	3.32	86.4	15.8	20.0
subject 4	3.33	78.5	11.2	14.1
average	3.33	82.9	13.7 ± 2.1	17.3 ± 2.7

applications such as defending against shoulder-surfing attacks, and provide assistive care for people with motor disability as well. The cheap and portable devices in our BCI system impose several challenges and complicate the entire system design. Through extensive experiments, we have designed an optimization framework to process the weak signal and have determined an optimal stimulus configuration by considering various factors.

We have built a prototype and evaluated it with real experiments. The bit rate of our system on computer monitor is 54.8 bits/minute on average, which is comparable to that of existing systems with expensive and laboratory-oriented devices and is the best so far with such portable devices. We believe that this is an appealing accomplishment in this area. Furthermore, we are the first to design and evaluate the SSVEP-based BCI system by displaying stimuli on a smartphone screen, with an average bit rate of 17.3 bits/minute.

References

[1] Philip Low and Stephen W. Hawking: Towards establishing neural correlates of intended movements and speech, http://www.neurovigil.com.

[2] Biosemi ActiveTwo EEG system, http://www.biosemi.com.

[3] Emotiv Epoc Headset, http://www.emotiv.com.

[4] G. Hackmann, W. Guo, G. Yan, Z. Sun, C. Lu, S. Dyke, Cyber-physical codesign of distributed structural health monitoring with wireless sensor networks, IEEE Transactions on Parallel and Distributed Systems (2014) 63–72.

[5] M. Keally, G. Zhou, G. Xing, J. Wu, A. Pyles, PBN: towards practical activity recognition using smartphone-based body sensor networks, in: SenSys, ACM, 2011, pp. 246–259.

[6] R. K. Ganti, P. Jayachandran, T. F. Abdelzaher, J. A. Stankovic, SATIRE: a software architecture for smart attire, in: Mobisys, ACM, 2006, pp. 110–123.

[7] J. Ko, T. Gao, A. Terzis, Empirical study of a medical sensor application in an urban emergency department, in: Proceedings of the Fourth International Conference on Body Area Networks, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2009, p. 10.

[8] G. Bin, X. Gao, Z. Yan, B. Hong, S. Gao, An online multi-channel SSVEP-based brain–computer interface using a canonical correlation analysis method, Journal of neural engineering 6 (4) (2009) 046002.

[9] I. Volosyak, SSVEP-based Bremen–BCI interfaceboosting information transfer rates, Journal of neural engineering 8 (3) (2011) 036020.

[10] NeuroSky MindWave, http://www.neurosky.com.

[11] A. Campbell, T. Choudhury, S. Hu, H. Lu, M. K. Mukerjee, M. Rabbi, R. D. Raizada, Neurophone: brain-mobile phone interface using a wireless eeg headset, in: Proceedings of the second ACM SIGCOMM workshop on Networking, systems, and applications on mobile handhelds, ACM, 2010, pp. 3–8.

[12] I. Martinovic, D. Davies, M. Frank, D. Perito, T. Ros, D. Song, On the feasibility of side-channel attacks with brain-computer interfaces, in: 21st USENIX Security Symposium, 2012, pp. 143–158.

[13] Y.-T. Wang, Y. Wang, T.-P. Jung, A cell-phone-based brain–computer interface for communication in daily life, Journal of neural engineering 8 (2) (2011) 025018.

[14] Y. Liu, X. Jiang, T. Cao, F. Wan, P. Mak, P. Mak, M. Vai, Implementation of SSVEP based BCI with Emotiv EPOC, in: IEEE VECIMS, IEEE, 2012, pp. 34–37.

[15] X. Zhang, I. S. MacKenzie, Evaluating eye tracking with iso 9241-part 9, in: HCI Intelligent Multimodal Interaction Environments, Springer, 2007, pp. 779–788.

[16] Z. Lin, C. Zhang, W. Wu, X. Gao, Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs, IEEE Transactions on Biomedical Engineering 53 (12) (2006) 2610–2614.

[17] J. R. Wolpaw, H. Ramoser, D. J. McFarland, G. Pfurtscheller, Eeg-based communication: improved accuracy by response verification, Rehabilitation Engineering, IEEE Transactions on 6 (3) (1998) 326–333.