Brain–Machine Interfacing Based on Visual Stimuli

PH.D. THESIS

KAORI SUEFUSA

A thesis submitted in fulfillment of the requirements

for the degree of of Doctor of Philosophy

Department of Electronics and Information Engineering

Tokyo University of Agriculture and Technology

February 12, 2019

Abstract

Decoding brain activities from brain signals is an important and challenging technology. An interface based on brain activities is called a brain machine/computer interface (BMI/BCI). BMIs capture brain activities evoked by mental tasks or external stimuli and translate them into commands for controlling devices or messages for users. This study is especially focused on BMIs with electroencephalogram (EEG) using repetitive visual stimuli, which elicit steady-state visual evoked potentials (SSVEPs).

A typical way of achieving command input with SSVEP-based BMI is subject to eye-movement. In a similar way, the command input can be implemented with an eye tracking system based on dwelling time. Evaluations of both interfaces have been widely reported in their own research communities, and it is acknowledged that both can accomplish high precision in terms of recognition accuracy. However, to our knowledge there have been no attempts to quantitatively compare their performances on the same experimental platform. Thus, firstly in this study, we compare between SSVEP-based BMI and dwelling-based ETI, and attempted to clarify their drawbacks and advantages. We evaluate their performance by investigating their accuracies and information transfer rate (ITR) with respect to the target size and the command analysis time, i.e., time window length of EEG analysis or the dwell time.

Among SSVEP-BMIs, a state-of-the-art is the one based on hybrid frequency and phase coded SSVEP, which needs a large set of calibration data as reference signals, called individual templates. This calibration can be laborious and time-consuming. Therefore, we propose a new approach for

reducing calibration time by transferring the individual template of a certain command, called a source template, to obtain new templates of the other commands, called target templates. The target templates are generated by shifting the frequency and phase of the source template to the desired frequency and phase. Moreover, we propose a method employing multiset canonical correlation analysis (MCCA) for detecting hybrid coded SSVEPs which aimed to compensate the decrease of recognition accuracy when calibration is reduced.

Equally importantly, the operations of the BMIs should be asynchronous in a practical use. However, a very limited number of works on asynchronous SSVEP-based BMIs have been reported. Most of the current BMIs based on SSVEP are synchronous, as the timing of a command entry is often indicated by some cues and is controlled by the BMIs. Moreover, the previous studies which proposed asynchronous SSVEP-based BMIs utilize only frequencies for visual stimuli, which led to a limited number of commands. Hence, we propose hybrid-coded visual stimuli in an asynchronous BMI, yielding an increase of the number of commands.

The experiments are conducted with 10–11 healthy subjects, and their EEG signals were analyzed offline. The experimental results showed that the BMI was comparable to the eye tracking system in terms of accuracy and the information transfer rate. In particular, when the size of a target was relatively small, the BMI had significantly better performance than the ETI. In addition, the proposed method for calibration reduction enabled to recognize both frequency and phase with a small amount of calibration closely achieving the performance using the full calibration data. Besides, the proposed method using MCCA showed a significant improvement in command recognition accuracy and ITR with the proposed method compared to the state-of-the-art. Finally, our proposed asynchronous BMI based on hybrid-coded SSVEP showed a promising performance and enabled a large increase

in the number of commands.

Declaration of Authorship

I, KAORI SUEFUSA, declare that the thesis I am submitting titled, "Brain–Machine Interfacing Based on Visual Stimuli" is entirely my original work. Some parts of this thesis including methods, texts, figures, tables etc. has previously been submitted for a degree, which have been used at these thesis. I have clearly stated the contribution of other authors that I have included in my thesis.

Signed:

_____

Date:

_____

Acknowledgements

First and foremost I would like to express my appreciation and gratitude to my advisor, Prof. Toshihisa Tanaka. His support and guidance have make the completion of this thesis possible. His guidance with patience and knowledge is the ultimate outcome of this thesis. Without his generous help it would have been impossible to finish this work.

I am also grateful to my family and all of my friends. They have always been supportive and I never forget that.

Lastly, I would like to express my appreciation to Mr. Hirohiko Araki, and his creation, Jojo's Bizarre Adventure. His work guided me whenever I almost lost my way. I am deeply grateful to his work.

This work is supported in part by JSPS Research Fellowship 16J08755.

Contents

List of Figures

List of Tables

Dedicated to JJBA...

Chapter 1

Introduction

1.1 Background

Brain–computer/machine interfacing (BMI/BMI) is an emerging potential application of signal processing, machine learning, and neuroscience that provides a communication pathway between a brain and an external world, enabling control of assistive technologies, environmental control, interaction with other people [1, 2, 3, 4, 5, 6, 7]. BMIs capture brain activities evoked by mental tasks or external stimuli and translate them into commands for controlling devices or messages for users. Thus, the BMIs enable people with severe motor disabilities to communicate and control devices without muscular movements. In addition, using feedback that people receive while observing the message, the BMIs can also be a tool for rehabilitation by inducing brain plasticity to restore motor function [8, 9].

Such brain activities for BMIs can be measured by noninvasive methods including electroencephalography (EEG) [10, 11]. In terms of uses for BMIs, the widely-used features can be divided into two types: exogenous and endogenous ones. Exogenous responses include event-related potentials (ERPs) evoked by sensory events such as visual, auditory, and somatosensory stimuli [12, 13, 14, 15]. These are also known as visual evoked potentials (VEPs) [16], auditory evoked potentials (AEPs) [17], and somatosensory evoked potentials (SSEPs) [15] depending on the type of associated sense.

VEPs elicited with repetitive visual stimuli are called steady state visual evoked potentials (SSVEPs) [18, 19], which are phase-locked responses to a periodic visual stimulus with a constant frequency [20], for example, flickering lights, images, checkerboards [19]. SSVEPs contain stationary periodic oscillations with the same frequency as the visual stimulus, its harmonics, and its subharmonics [21].

Endogenous responses include ERPs evoked by cognitive or motor events, and event-related (de)synchronization (ERD/ERS) also evoked by cognitive or motor events. Cognitive events are, for example, ones related to attention, memory, and language. Motor events are, for example, ones related to the preparation or the imagery of movement [22, 23, 24]. Generally speaking, BMIs based on exogenous responses require less user training than that based on endogenous responses [25, 19]. Specifically, BMIs based on SSVEPs have high recognition accuracies and require only a short time to enter a command; hence, they result in a high information transfer rate (ITR) [26], which is a measure used to evaluate the accuracy, the mean time to enter one command, and the number of choices/commands [2].

1.2 Problems

A typical way of achieving command input with SSVEP-based BMI is as follows. First, multiple visual stimuli whose flickering frequencies differ from one another are shown on a screen to the user. Each visual stimulus corresponds to each command. Next, the user gazes at one of the visual stimuli that corresponds to his or her desired command. Then, with proper signal processing, the BMI recognizes which stimulus the user is gazing at, that is, which command the user would like to input. Conversely, the target on a screen at which the user gazes can also be measured by tracking the point of the gaze, that is, where the user is looking. The technique that tracks the

point of the gaze is called eye tracking, and interfaces that allow command input using eye tracking are often called eye tracking interfaces (ETIs) [27, 28, 29, 30]. Among several techniques for command entry on ETIs, inputting a command by gazing at a target for more than predefined time—known as dwelling—is one of the most straightforward and commonly used methods [31], and it does not require any physical movement of the limbs. Both BMI and ETI are user interfaces that utilize visual perception. Evaluations of both interfaces have been widely reported in their own research communities [32, 33], and it is acknowledged that both can accomplish high precision in terms of recognition accuracy [34, 35]. However, to our knowledge there have been no attempts to quantitatively compare their performances on the same experimental platform.

Among SSVEP-BMIs, a state-of-the-art is the one based on hybrid frequency and phase coded SSVEP, which needs a large set of calibration data as reference signals, so-called individual templates. A major issue which still remains on these algorithms is that they require a long time for calibration to obtain an acceptable individual template. Such a calibration step can take a lot of time and energy and might cause a visual fatigue before an actual use.

A method for calibration reduction has been proposed by authors [36]. It enabled massive reduction of calibration time, however, it also led to a cutback of recognition accuracy. Recently, a method of detecting hybrid-coded SSVEPs has been proposed [37]. The method employed canonical correlation analysis (CCA)[38, 39] utilizing a training set, and enabled increasing a number of commands, improving a recognition accuracy, and boosting an information transfer rate (ITR) [2]. However, this CCA-based method requires to solve a generalized eigenvalue problem several times to obtain multiple spatial filters which enhance SSVEP components, and there is a room for improvement in the performance yet.

Additionally, SSVEP-based BMIs are advantageous in terms of the recognition accuracy and the ITR, most of their operations are required to be *asynchronous* in a practical use [40, 41, 42]. However, a very limited number of works on asynchronous SSVEP-based BMIs have been reported [43, 44, 41, 45, 46]. Most of the current BMIs based on SSVEP are *synchronous*, as the timing of a command entry is often indicated by some cues and is controlled by the BMIs, and operations are executed in a certain time interval. In other words, synchronous BMIs attempt to interpret all signals which a brain constantly generates as a user intent.

The study focuses on four following unexplored problems and challenges regarding to SSVEP-based BMI;

1) comparison with ETI

2) calibration reduction

3) achieving higher recognition accuracy

4) asynchronousness

1.3 Solutions

We propose a method to solve the problems 1), 2), 3), and 4) in Chapter 3, 4, 5, and 6, respectively. In Chapter 3, we compare between BMI and ETI, and clarify their drawbacks and advantages. Specifically, we compare between a SSVEP-based BMI and a dwelling-based ETI. We evaluate their performance by investigating their accuracies and ITR with respect to the target size and the command analysis time, i.e., time window length of EEG analysis or the dwell time.

Chapter 4 utterly proposes a novel approach to calibration reduction by transferring the individual template of a certain command, called a source

template, to obtain new templates of the other commands, called target templates. The target templates are generated by shifting the frequency and phase of the source template to the desired frequency and phase. An experiment with a mixed-coded SSVEP-BMI was conducted to evaluate performance of the proposed method in terms of the recognition accuracy.

In Chapter 5, we propose to employ multiset canonical correlation analysis (MCCA) [47] for detecting mixed coded SSVEPs in order to simplify the CCA-based method [37]. Since MCCA enables us to evaluate a relationship between multiple datasets simultaneously, the proposed method requires to solve the generalized eigenvalue problem just once in order to obtain spatial filters. Moreover, we hypothesize that we can expect higher recognition accuracy by utilizing the eigenvalue as a feature. An experiment was conducted to evaluate performance of the proposed method in terms of the recognition accuracy and the ITR.

In Chapter 6, we propose to use not only frequency but also phase, known as mixed-coded visual stimuli [48, 37, 49] in an asynchronous BMI, yielding an increase of the number of choices/commands. So as to detect the phase of SSVEPs asynchronously, the visual stimuli were presented as flickers putting a non-flickering blank interval between each flickering interval to synchronize the recorder of EEG with the stimuli. For discriminating between IC and NC states as well as detecting commands in the case of mixed-coded visual stimuli, the proposed system incorporates a novel method that exploits MCCA [47, 49], yielding spatial filters that enhance the SSVEP component. We hypothesized that the proposed implementation with the proposed MCCA decoder would increase the accuracy in command detection compared to the conventional decoders in the classical setting of frequency-coded stimuli.

1.4 Organization of Thesis

As depicted in Figure 1.1, this thesis is divided into seven chapters. In Chapter 1, the background regarding this work, the problems, the ideas of the proposed method is discussed. Chapter 2 introduces the fundamentals of BMIs. Chapter 3 describes the comparison between BMI and ETI. In Chapter 4, we propose a novel approach to calibration reduction. In Chapter 5, a method for detecting mixed coded SSVEPs is introduced. In Chapter 6, a novel mixed-coded SSVEP-based BMI is proposed. Finally, Chapter 7 concludes the thesis.

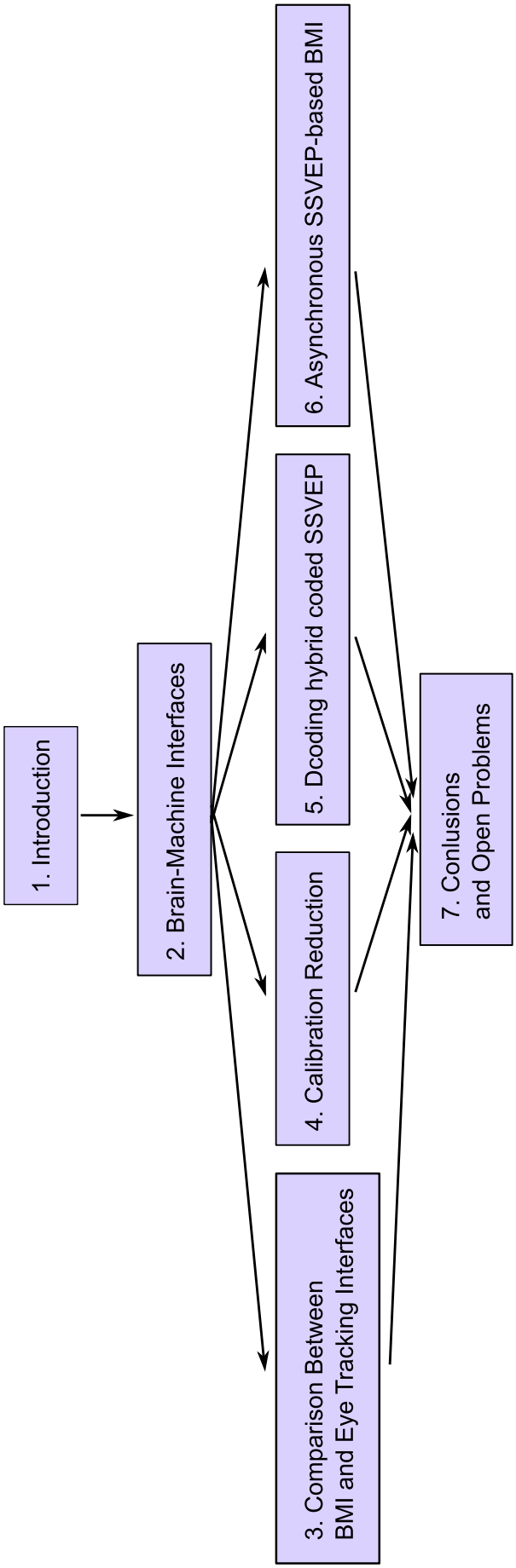


FIGURE 1.1: Diagram for the organization of the thesis.

Chapter 2

Brain–Machine Interfaces

In this chapter, the overview of BMIs is described. In Section 2.1, we describe fundamental structures of BMI. In Section 2.2, several methods for measuring brain activities are listed. In Section 2.3, some EEG features used in BMIs are described. Finally in Section 2.4, we briefly introduce asynchronous BMI.

2.1 Fundamentals of BMI

BMI provides a communication pathway between a brain and an external world [1, 2, 5, 6, 7]. In terms of "the brain to the external world", BMIs capture brain activities evoked by mental tasks or external stimuli and enable people with severe motor disabilities to communicate and control devices without muscular movements. Thus, the BMIs enable people with severe motor disabilities to communicate and control devices without muscular movements. In addition, using feedback which people receives observing the output, the BMIs can also be a tool for rehabilitation by inducing brain plasticity to restore motor function [8, 9].

2.1.1 Scopes

Therefore, BMI has a potential to help people with severe motor disabilities, such as amyotrophic lateral sclerosis (ALS) [50]. The number of people living

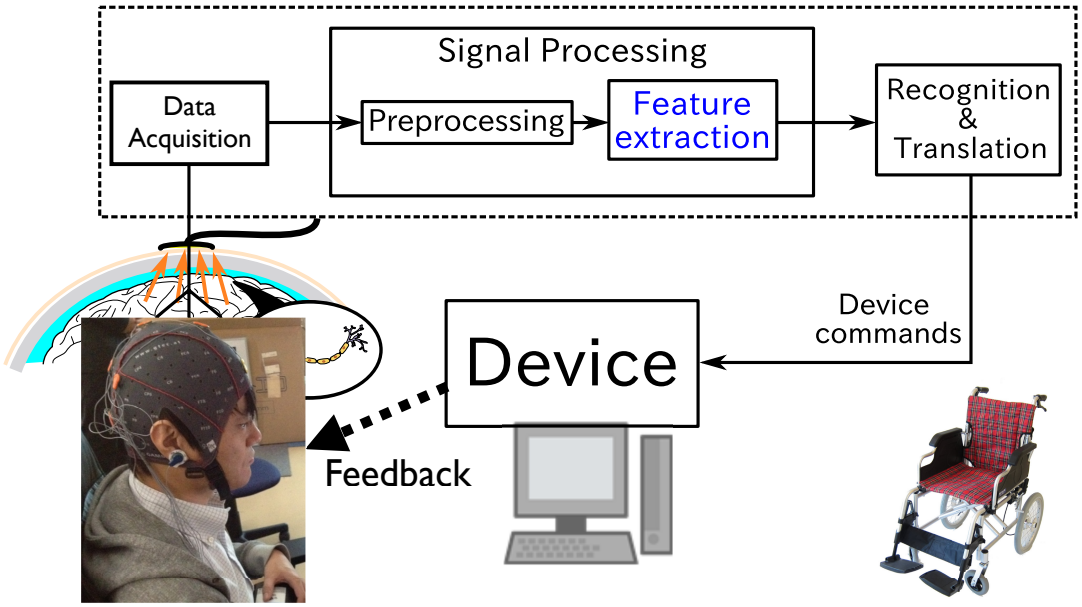


FIGURE 2.1: Structure of BMI.

with ALS is estimated at around 450,000 worldwide. Additionally, global average life expectancy have been increasing for last several decades, which leads to the aging society. BMI can be one of the way to assist those aging and disable population since it does not require any muscular movements of limb.

2.1.2 Design

Figure 2.1 shows an essential structure of BMI. As shown in the figure, a procedure of BMI system can be divided into the following four steps.

Step 1 A user performs a certain task which induces a specific neuronal activation. The task can be mental tasks such as imagination of something or turning attention to a visual, auditory, tactile stimulus, and so forth.

Step 2 Brain activities are measured by a measurement system. The acquired signals are converted to the digital signal by an A/D converter after amplification and filtering are applied.

Step 3 Signal processing is applied for preprocessing and feature extraction. First, recorded signal is preprocessed to reduce noise and make the features clear for detection. Next, the feature components related to the certain task are extracted since the signal is mixed with components associated with various brain activities.

Step 4 The task performed by the user is estimated by recognizing the extracted features. The classification algorithms with machine learning including linear discriminant analysis (LDA), support vector machine (SVM), artificial neural network (ANN) are widely used.

Step 5 The estimated task are translated into a command entry or feedback. The command entry can be used to operate a machine, and the feedback can be used for rehabilitation.

2.2 Measuring brain activities

2.2.1 Invasive methods

Methods for measuring brain activities for BMI can be coarsely classified as invasive and noninvasive. Invasive ways of measuring of measuring brain activities for the BMIs include Electrocorticogram (ECoG), electrosubcorticogram (ESCoG), and electroventriculogram (EVG) They require surgeries to install electrodes on a cortex or a cerebral ventricle and measure brain activities. Invasive ways can measure the brain activities with much less noise than noninvasive ways, whereas, they have the risks generated by permanent surgically implanted devices in the brain.

2.2.2 Noninvasive methods

Noninvasive ways are such as electroencephalography (EEG), magnetoencephalography (MEG), near-infrared spectroscopy (NIRS), and functional magnetic resonance imaging (fMRI) [10]. Each technique and their pros and cons are described below.

Electroencephalography

EEG is a technique which measures brain activities with electrodes. The electrical activity which can be observed by an electrode is the summation of activities of the group of neurons near the electrode. The measurement method which locates the electrodes on a scalp, which means it is noninvasive, is called scalp EEG. Compared to invasive methods, scalp EEG has relatively low SNR. In the remaining of the thesis, we call the scalp EEG simply EEG unless otherwise specified. The equipment of EEG is more compact compared to other systems such as MEG and fMRI. In addition, its temporal resolution is also higher than them, however, its spatial resolution is lower because of the size of the electrodes

Magnetoencephalography

MEG is an imaging technique which employs superconducting quantum interference device (SQUID). SQUID measures magnetic field generated by electrical currents occurring naturally in the brain. It has high temporal and spatial resolution, although it also has disadvantages such as its high cost and large equipment size.

Near-infrared spectroscopy

NIRS uses the near-infrared spectrum which passes through a scalp and a skull and reaches the inside. It measures the reflected light and the change

of amount of oxyhemoglobin and deoxyhemoglobin, which are reflected by the exchange of oxygen in the brain. Since the hemodynamic response is relatively slow, its temporal resolution is limited.

Functional magnetic resonance imaging

FMRI is a technique which uses magnetic resonance imaging to measure brain activities. When neuron activates in one area of the brain, blood flow increases and oxygen is consumed there. FMRI measures brain activities by detecting associated changes in the blood flow. It has a high spatial resolution and can capture activities from deep parts of the brain, whereas its temporal resolution is low since the changes of blood flow is slower than neuron activations.

Summary

It is considered that the EEG is most practical for BMI applications since it has high temporal resolution and ease of use, moreover, its hardware cost is relatively lower than other techniques [10]. Although the EEG signal has a small amplitude in the range of microvolts and its signal-to-noise ratio (SNR) is low because of environmental noise, biological noise, and artifacts [11]. Necessarily, proper signal processing is required to extract useful features contained in the EEG [51].

2.3 EEG-based BMI

In terms of the use for the BMIs, widely-used features which appear in the EEG can be divide into two types of exogenous and endogenous ones. Exogenous responses include event-related potentials (ERPs) evoked by sensory events such as visual, auditory, and somatosensory stimuli [12, 13, 14,

15]. These are also known as visual evoked potentials (VEPs) [16], auditory evoked potentials (AEPs) [17], and somatosensory evoked potentials (SSEPs) [15] depending on the type of the associated sense. AEPs elicited with rapid auditory stimuli is called auditory steady state responses (ASSRs) [14, 19], and VEPs elicited with rapid visual stimuli is called steady state visual evoked potentials (SSVEPs) [18, 19]. Endogenous responses include ERPs evoked by cognitive or motor events, and event-related (de)synchronization (ERD/ERS) also evoked by cognitive or motor events. Cognitive events are, e.g., ones related to attention, memory, and language. Motor events are, e.g., ones related to preparation or imagery of movement [22, 23, 24]. Compared to BMIs based on endogenous responses, BMIs based on exogenous responses require less training of a user [25, 19].

2.4 Synchronous/Asynchronous BMI

The operations of the BMIs are required to be *asynchronous* in a practical use [40, 41, 42]. However, a very limited number of works on asynchronous BMIs have been reported [43, 44, 41, 45, 46]. Most of the current BMIs are *synchronous*, as the timing of a command entry is often indicated by some cues and is controlled by the BMIs, and operations are executed in a certain time interval. In other words, synchronous BMIs attempt to interpret all signals which a brain constantly generates as a user intent. The issue is often referred to as "Midas touch problem" [3, 52]. However, in a real environment, the user does not intend to enter commands for a certain period. In other words, a practical BMI should be asynchronous [40, 41, 42]. Such a BMI is known as an asynchronous or a self-paced BMI, which enables users to enter commands only when they intend to do so. To implement asynchronous BMIs, a state when the user does not intend to enter commands needs to be considered. Such a state is typically called a non-control (NC) or idle state,

while a state when the user actually intends to enter commands is called an intentional control (IC) or work state [53, 42, 54]. Thus, a key issue to implement asynchronous BMIs is to establish an efficient decoder for identifying not only the states (NC/IC) but also the command.

Chapter 3

Comparison Between BMI and Eye Tracking Interface

BMI based on visual stimuli detects the target on a screen on which a user is focusing. The detection of the gazing target can be achieved by tracking gaze positions with a video camera, which is called eye tracking or eye tracking interfaces (ETIs). Both types of interfaces have been developed in different communities. Thus, little work on the comprehensive comparison between these two types of interfaces has been reported. In our study, we quantitatively compare the performance of these two interfaces on the same experimental platform. Specifically, our study is focused on two major paradigms of BMI and ETI: steady-state visual evoked potential-based BMIs and dwelling-based ETIs. Recognition accuracy and the information transfer rate were measured by giving subjects the task of selecting one of four targets by gazing at it. The targets were displayed in three different sizes (with sides 20, 40, and 60 mm long) to evaluate performance with respect to the target size. The experimental results showed that the BMI was comparable to the ETI in terms of accuracy and the information transfer rate. In particular, when the size of a target was relatively small, the BMI had significantly better performance than the ETI. The results on which of the two interfaces work better in different situations would not only enable us to improve the

design of the interfaces but would also allow for the appropriate choice of interface based on the situation. Specifically, one can choose an interface based on the size of the screen that displays the targets.

3.1 Introduction

A typical way of achieving command input with SSVEP-based BMI is as follows. First, multiple visual stimuli whose flickering frequencies differ from one another are shown on a screen to the user. Each visual stimulus corresponds to each command. Next, the user gazes at one of the visual stimuli that corresponds to his or her desired command. Then, with proper signal processing, the BMI recognizes which stimulus the user is gazing at, that is, which command the user would like to input. For the visual stimulus, a variety of visual reversal patterns, such as square box and checkerboard patterns, are used [55]. Flickering stimuli can be strenuous, causing visual fatigue; there is also a risk of triggering epileptic seizures in certain people [33, 56].

Conversely, the target on a screen at which the user gazes can also be measured by tracking the point of the gaze, that is, where the user is looking. The technique that tracks the point of the gaze is called eye tracking, and interfaces that allow command input using eye tracking are often called eye tracking interfaces (ETIs) [27, 28, 29, 30]. They are considered to be fairly robust [35] and promising components of user interfaces [57] because of their inherent advantages, such as their speed and ease of use [35]. ETIs allow for performing actions or selections as well as pointing on a screen. Methods that enable command input include selecting a target by gazing or dwelling for more than a fixed time [35], using a physical button [35], blinking or winking, making gestures with the gaze [58], or using eye movements with manual

pointing, which is called manual and gaze input cascaded (MAGIC) pointing [59]. Among them, inputting a command by gazing at a target for more than predefined time—known as dwelling—is one of the most straightforward and commonly used methods [31], and it does not require any physical movement of the limbs. In most cases, it is easy to mount an eye tracking device called an eye tracker. However, it requires the user to calibrate the device prior to use. A little movement of the eye position can deteriorate the performance that necessitates recalibration. In addition, targets need to be a certain size to compensate for inaccuracies in the estimation of the gaze coordinates [60].

Both BMI and ETI are user interfaces that utilize visual perception. Evaluations of both interfaces have been widely reported in their own research communities [32, 33], and it is acknowledged that both can accomplish high precision in terms of recognition accuracy [34, 35]. A previous study on a SSVEP-based BMI implemented as a speller [61] claimed that its typing speed up to 12 words per minute was comparable to the reported typing speed of typical ETIs, typically from 5 to 10 words per minute [62]. Another study investigated and compared a participant's performance and feedback using an auditory ERP-based BMI, an ETI, and electrooculography [63]. The participant rated the BMI as the easiest to use and the ETI as the least tiring, while another comparison study on a visual ERP-based BMI and an ETI reported that the ETI was superior to the BMI in terms of their performances and usability [64]. Meanwhile, the preliminary work of the authors compared the performances of a SSVEP-based BMI and a dwelling-based ETI by investigating recognition accuracies and information transfer rates (ITRs) [65] in which subjects were asked to gaze at one of four targets on a screen [66]. The experimental results showed that the BMI was comparable in ITR to the ETI and had better performance for the relatively small targets on the screen. However, the preliminary work included a small number of subjects without

statistical validation. Thus, the purpose of our study is to investigate and compare the performances of the SSVEP-based BMI and the dwelling-based ETI in detail, employing more subjects and more detailed statistical analyses. As mentioned above, these interfaces have been evaluated in their own research communities regarding accuracy, ITR, usability, and so on. To avoid broadening the issue too much, the present study is focused on investigating performance, including recognition accuracy and ITRs, not usability, including visual fatigues and workloads. Their performances were evaluated with respect to the size of the targets and the time to analyze a command, that is, the time window length of EEG analysis in the BMI and the dwell time in the ETI.

3.2 Methods

3.2.1 Subjects

Ten subjects (two females, age range = 22–27, mean age = 23.4 $\pm$ 1.6) participated in the experiment. All subjects were healthy and had normal or corrected-to-normal vision. They were given an informed consent form, and the study was approved by the Research Ethics Committee of Tokyo University of Agriculture and Technology.

3.2.2 Experimental settings

BMI

We used Ag/AgCl active electrodes manufactured by Guger Technologies (g.tec), namely, g.LADYbird, g.LADYbirdGND (for GND), and g.GAMMAearclip (for reference, earclip type) to record EEG data. These were driven by a g.GAMMAbox (g.tec) power supply unit. As illustrated in figure 3.1, four electrodes were located at Pz, Oz, O1, and O2 following the international

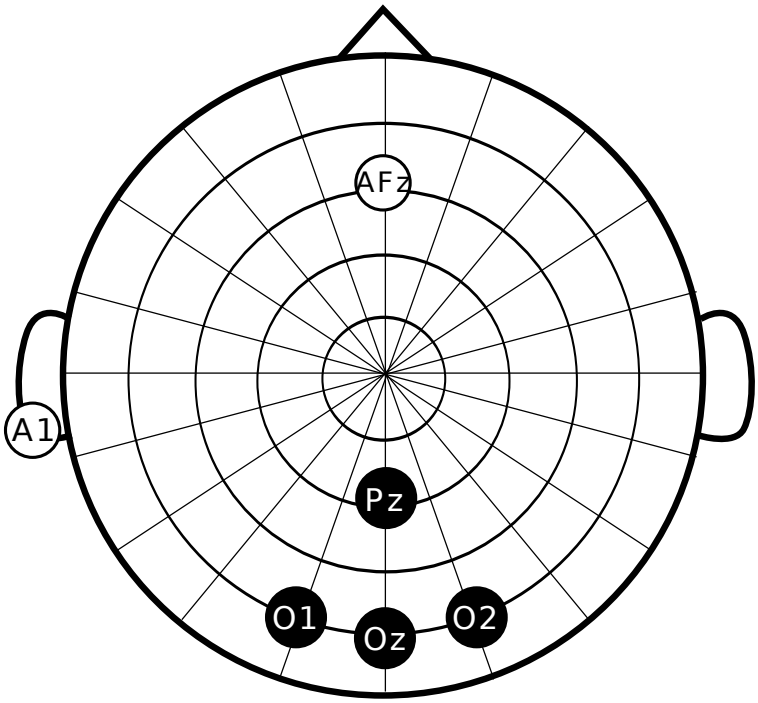


FIGURE 3.1: Electrode positions.

10–20 system covering the occipital region. The electrodes for GND and the reference were placed on AFz and A1, respectively. The EEG signals were amplified by MEG-6116 (Nihon Kohden), which applied lowpass and high-pass analog filters for each channel. The cutoff frequencies of the lowpass and highpass filters were set to 100 Hz and 0.5 Hz, respectively. The signals were sampled by an A/D converter (AIO-163202F-PE, Contec) with a sampling rate of 1200 Hz and recorded and downsampled to 240 Hz with Data Acquisition Toolbox of the MATLAB (MathWorks).

ETI

EyeFrame SceneCamera System (Arrington Research) was used as an eye tracker, which can be worn as frames of glasses. The system has an approximate measurement of accuracy of 0.25–1.0 degree of visual angle [67]. The dark pupil method [68] was used to determine the position of the pupil. With this method, an infrared light source was placed in front of the eye, and it

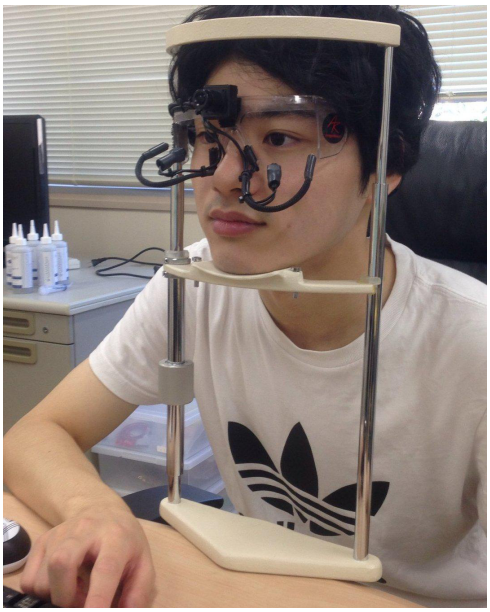


FIGURE 3.2: Subject using a chin rest.

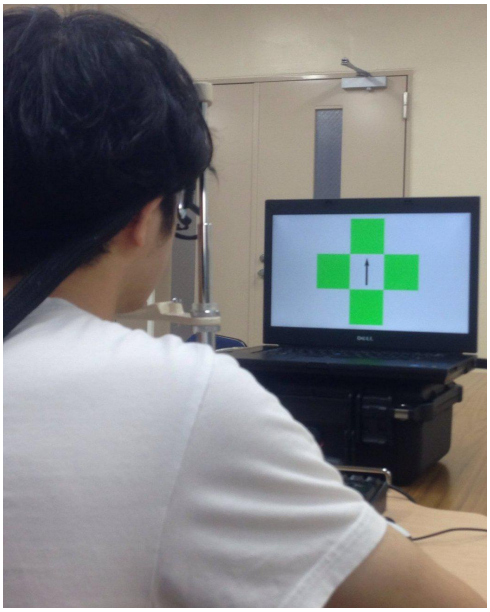
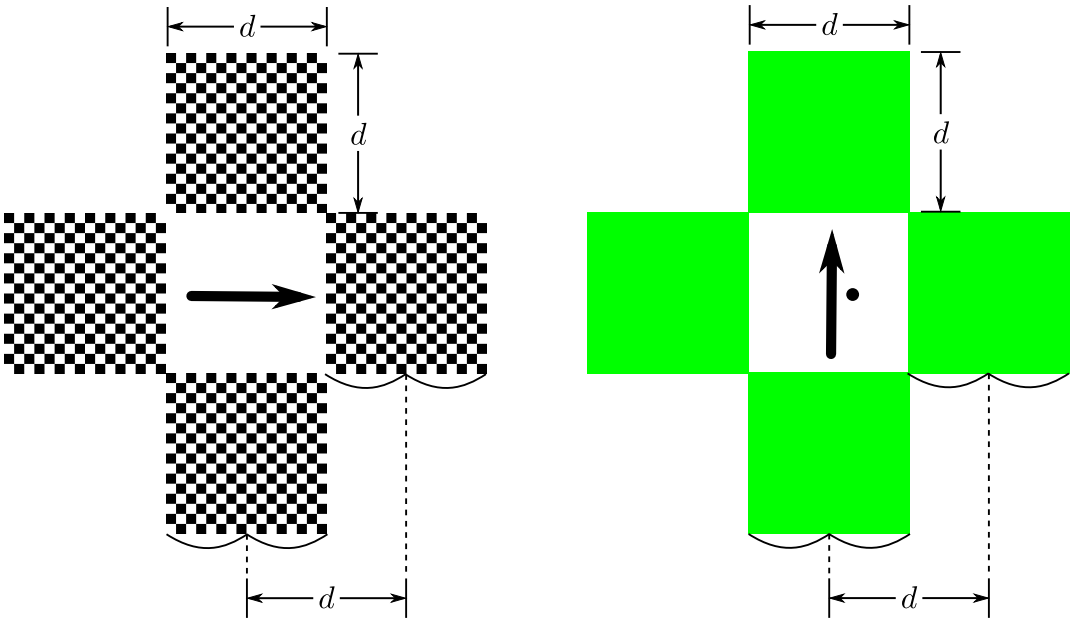


FIGURE 3.3: Experimental circumstance.

made the iris appear light, with the pupil being the darkest region in the image. Calibration was performed using Auto-Calibration of ViewPoint Eye-Tracker (Arrington Research), which was required to map the coordinates as the point of gaze on the screen. During calibration, the subjects were asked to fixate on a 3×3 grid of points that were displayed one at a time in random order. After calibration, the xy coordinates of the user's gaze point were smoothed by averaging previous four samples and recorded using MATLAB (MathWorks) with a sampling rate of 60 Hz. Throughout the experiment, the subjects' heads were fixed using a chin rest. Figure 3.2 shows a subject using the chin rest during an experiment. A MATLAB Psychtoolbox was used to plot a black dot with a diameter of 2 mm as the feedback of the subject's gaze point, as depicted in figure 3.4b. Figure 3.3 shows the circumstance of the experiment.

TABLE 3.1: Flickering frequencies corresponding to visual targets on the screen, which were assigned to BMI commands.

Target T_k	Location	Frequency f_k
T_1	Top	9.2 Hz
T_2	Right	10.9 Hz
T_3	Bottom	12.0 Hz
T_4	Left	13.3 Hz



(A) BMI (B) ETI

FIGURE 3.4: Displayed targets. The black dot shown in (b) represents the gaze point of a user. Through the experiments, three different sizes (d = 20, 40, and 60 mm) per side were used for comparison.

3.2.3 Design

The targets shown in figure 3.4 were drawn with Psychtoolbox on a display screen. For the BMI experiment, we used a desktop computer connected to a 23-inch display with a resolution of 1920 $\times$ 1080 and a refresh rate of 120 Hz. As depicted in figure 3.4a, the targets in the BMI experiment were square checkerboards that reversed pattern according to the frequencies listed in table 3.1. The length of one side of each small square of the checkerboards was 4 mm (visual angle of 0.32 deg). For the ETI experiment, we used a

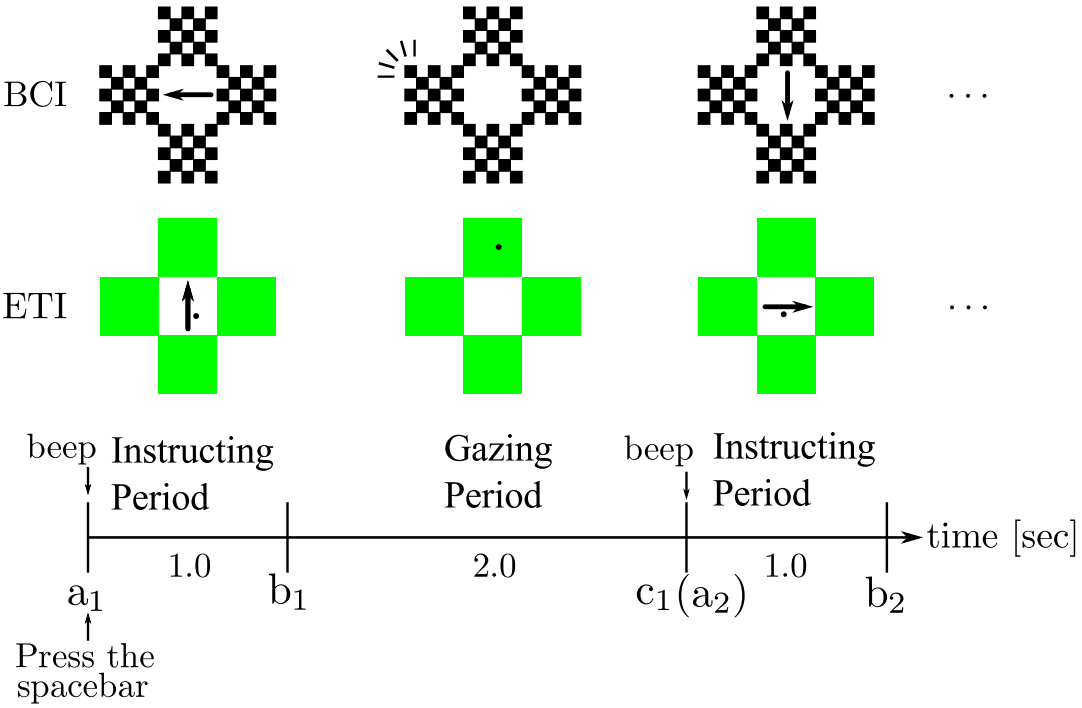


FIGURE 3.5: Schematic depiction of time course in one trial and an image of the displayed targets.

laptop computer connected to a 15.6-inch display with a resolution of 1366 $\times$ 768 and a refresh rate of 60 Hz. As can be seen in figure 3.4b, the targets in the ETI experiment were drawn in green so that they did not match up with the colors of the background or the gaze-point which were drawn in gray and black, respectively. In both experiments, subjects sat on a comfortable chair in front of the display 70 cm away, and the display was set at a height they could look straight.

As illustrated in figure 3.4, four square targets were displayed on the top, bottom, right, and left on the screen. The targets were referred to as T_k in clockwise order from the top; thus, a target on top was T_1 and a target on the left was T_4, as shown in table 3.1. The three different target sizes employed in the experiments were 20, 40, and 60 mm on a side, which can be converted to visual angles of 1.6, 3.3, and 4.9 deg. The distance between the center of the targets and each target was the same as the target size, In other words, the distances between two adjacent targets was proportional to the target size.

As depicted in figure 3.4, an arrow was also displayed in the center of the targets to dictate which target the subject should gaze at.

3.2.4 Task

Each trial consisted of one second of an instructing period and two seconds of a gazing period, as depicted in figure 3.5. The instructing period (e.g., the interval of a_1–b_1 in figure 3.5) started with a beep sound for a duration of 0.1 seconds. During the instructing period, the arrow was presented to instruct the subjects in terms of which target to gaze at. After that, the gazing period (e.g., the interval of b_1–c_1 in figure 3.5) started with the disappearance of the arrow. The subjects fixated on the arrow during the instructing period, and following the disappearance of the arrow, they started to gaze at the target. The instructions corresponding to the targets were given equally and randomly.

As described before, three different target sizes (d = 20, 40, and 60 mm) were used for comparison. For each target size, four sessions consisting of five successive trials were executed. Therefore, twenty trials were executed for each target size, and one command was determined per trial. After each session, the task was temporary halted concerning the burden of the subjects' eyes. It was restarted when the subjects were ready and pressed the space bar.

As depicted in figure 3.5, in the BMI experiment, visual stimuli did not flicker during the instructing period, but flickered during the gazing period. Moreover, in the ETI experiment, calibration was performed before the experiment and after each of the four sessions because the subjects' heads or the eye tracker they wore could easily move out of position, which had a major effect on mapping the gaze point on the screen.

3.2.5 Target detection methods

The recordings (the EEG signals for the BMI and the gaze positions for the ETI) during the gazing period (the interval of b–c in figure 3.6) were analyzed offline, and entered commands were determined.

BMI

To detect commands, the EEG for W seconds during the gazing period was analyzed. Samples just after the visual stimuli started flicker were neglected and considered as a delay of SSVEP onset [69], as described later.

We employed a method called filter bank canonical correlation analysis (FBCCA), which was recently proposed by Chen *et al.* [70] for detecting the stimulus frequency without any calibration. The calibration-free method was employed despite the fact that methods that utilize calibration data have recently been gaining attention [71, 72, 73, 74, 75], since those methods sometimes take much more time for calibration than the ETI. FBCCA is an extension of the standard canonical correlation analysis (CCA) method [76], which finds the target frequency that maximizes the canonical correlation between the observed multichannel EEG and a vector consisting of simulated sinusoidal waves. More specifically, let $\boldsymbol{x}(t) \in \mathbb{R}^M$ be an M-channel EEG signal and $\boldsymbol{y}(t) \in \mathbb{R}^{2 \cdot H}$ consists of Fourier basis functions of simulated stimulus signals, which are ideal SSVEP with frequency f given as

$$\boldsymbol{y}_f(t) = \begin{bmatrix} \sin(2\pi f t) \\ \cos(2\pi f t) \\ \vdots \\ \sin(2H\pi f t) \\ \cos(2H\pi f t) \end{bmatrix}, \tag{3.1}$$

where H is the number of harmonics in the Fourier basis functions. To detect the frequencies of the SSVEP components contained in the EEG for the SSVEP-based BMI systems, a canonical correlation ρ_f between $\boldsymbol{x}(t)$ and $\boldsymbol{y}_f(t)$ is calculated as follows:

$$\rho_f = \max_{\boldsymbol{w}_{\boldsymbol{x}}, \boldsymbol{w}_{\boldsymbol{y}_f}} \frac{\boldsymbol{w}_{\boldsymbol{x}}^T E[\boldsymbol{x}(t)\boldsymbol{y}_f^T(t)]\boldsymbol{w}_{\boldsymbol{y}_f}}{\sqrt{\boldsymbol{w}_{\boldsymbol{x}}^T E[\boldsymbol{x}(t)\boldsymbol{x}^T(t)]\boldsymbol{w}_{\boldsymbol{x}} \boldsymbol{w}_{\boldsymbol{y}_f}^T E[\boldsymbol{y}_f(t)\boldsymbol{y}_f^T(t)]\boldsymbol{w}_{\boldsymbol{y}_f}}}, \tag{3.2}$$

which is calculated for all the frequencies of flickering targets. Thus the gazing target is identified by the frequency that gives the largest canonical correlation among the possible frequencies.

In the FBCCA, the observed EEG is decomposed into J subbands by a filter bank, and for the j-th ($j = 1, \ldots, J$) subband EEG signal, $\boldsymbol{x}_j(t)$, the canonical correlation with the reference signal, $\boldsymbol{y}_{f_k}(t)$, with the frequency of the k-th target, denoted by $\rho_{k,j}$, is calculated. Once the canonical correlation $\rho_{k,j}$ is calculated, the feature for f_k is defines as a weighted sum of squares of the correlation as follows:

$$\tilde{\rho}_k = \sum_{j=1}^{J} \boldsymbol{w}(j) \cdot \rho_{k,j}^2, \tag{3.3}$$

where the weight vector with respect to $\rho_{k,j}$ corresponding to subband components is suggested [70] to be

$$\boldsymbol{w}(j) = j^{-a} + b, \;\; j \in [1, J], \tag{3.4}$$

where a and b are constants. Thus, the frequency $f^\star$ corresponding to $k^\star$ that maximizes $\tilde{\rho}_k$ is detected as the stimulus frequency:

$$f^\star = f_{k^\star}, \tag{3.5}$$

TABLE 3.2: Subbands of the designed filter bank.

Subband$_j$	Passband [Hz]
SB$_1$	[9.2 66.5]
SB$_2$	[18.4 66.5]
SB$_3$	[27.6 66.5]
SB$_4$	[36.8 66.5]
SB$_5$	[46.0 66.5]

where

$$k^\star = \arg\max_k \tilde{\rho}_k. \tag{3.6}$$

In this chapter, the number of harmonics was set as $H = 5$. In additions, the subbands were designed so that they cover multiple harmonic frequency bands with individual low-cut and high-cut frequencies, as listed in table 3.2. As shown in table 3.2, the number of subbands was set as $J = 5$ empirically. See [70] for details on the filter bank design. Empirically, the constants in (3.4) were set as $a = 2$ and $b = 1$.

ETI

The operation of the dwelling-based ETI is typically self-paced, and a user can avoid involuntary command entry by simply looking away from a target before exceeding the dwell time, in other words, there is no input at times [77, 78, 31]. Conversely, the SSVEP-based BMI employed in the present study detects commands synchronously, and there would be some sort of input invariably after an analysis for target detection, as in most of the SSVEP-based BMI [34, 79, 61]. To fairly compare their performances, the entered commands were determined in two ways: 1) direct target selection (ETI-target) and 2) nearest target selection (ETI-nearest). The details of these analyses are described below.

ETI-target The target was detected as a command when the gaze point was dwelled on it directly for more than the dwell time D [sec]. If the gaze

point was outside of the target, the dwell timer was reset. In the case where the gaze point was not dwelled on any target for more than D [sec], the detection was regarded as a failure, and no command was detected. Thus, the detected result could be one of the four commands or no input.

ETI-nearest In the case where the gaze point did not satisfy the dwell time during the gazing period in accordance with the ETI-target approach, the target was determined based on the count of the nearest target of the gaze points during the gazing period. Thus, the detected result could be one of the four commands.

In terms of a four-command user interface, the ETI-nearest is more practical than the ETI-target achieving higher accuracy. However, it should be emphasized that, as the number of commands increase, the size of the targets becomes smaller and the targets are more densely located on the screen. In such a case, the ETI-nearest is less practical, as it is necessary for users to gaze directly at the target.

3.2.6 Performance criteria

To evaluate performance, the recognition accuracy and the ITR in bits/min [2] were calculated:

$$\mathrm{ITR} = \frac{60}{U}\left[\log_2 K + P\log_2 P + (1-P)\log_2 \frac{1-P}{K-1}\right], \tag{3.7}$$

where U is the mean time in seconds to input one command, K is the number of selectable commands ($K = 4$), and P is the recognition accuracy, that is, the probability that the detected commands match the instructed commands.

Furthermore, the simulated online performance was also investigated to evaluate the estimated performance in online use. It should be noted that

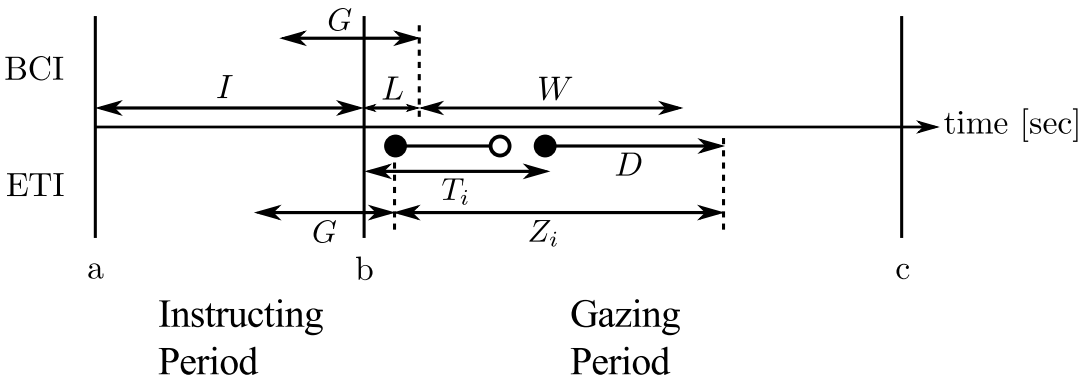


FIGURE 3.6: An example of the procedure for inputting a command. For the ETI (i.e., the lower part of the picture), a closed circle represents the second that the gaze point lands on the target, while an open circle represents the second it is off the target (a, b, and c correspond to those in figure 3.5).

the offline recognition accuracy was equal to the simulated online one since the same recorded data were analyzed for the evaluation of both the offline and the simulated online performance. In this chapter, the estimated ITR in online use considering the expected time to shift the gaze was defined as the *simulated online ITR*, while the ITR that was calculated considering the actual time required to input the commands, including breaks (the instructing periods), was defined as the *offline ITR*. The definition of U in (3.7) was altered with respect to the offline and the simulated online ITR individually, as described below. The simulated online ITR was estimated only for the analysis time, which led to the highest performance in the offline evaluation.

BMI

The ITR of the BMI was calculated with respect to two independent variables, the size of targets (d [mm]) and the time to analyze a command, that is, the data length of EEG signal analysis (W [sec]). To calculate the offline ITR, U in (3.7) was defined as follows:

$$U_{\mathrm{BMI}}^{\mathrm{off}} = I + L + W, \tag{3.8}$$

where I [sec] is the instructing period and L [sec] is the time lag (the delay) of SSVEP onset, as illustrated in figure 3.6. We set $I = 1.0$ sec, $L = 0.1$ sec, and $W = 0.5, 0.6, \ldots, 1.8, 1.9$ sec.

Additionally, to calculate the simulated online ITR, U in (3.7) was defined as follows:

$$U_{\mathrm{BMI}}^{\mathrm{on}} = G + W, \tag{3.9}$$

where G [sec] is the time for shifting the gaze from one target to another, as illustrated in figure 3.6. In the definition of U for calculating the simulated online ITR, the time lag of SSVEP onset L was removed, since it does not affect online use. We set $G = 0.5$ sec according to the previous studies [80, 61].

ETI

The ITRs of the ETI-target and the ETI-nearest were also calculated with respect to two independent variables, the size of targets (d [mm]) and the time to analyze a command, i.e., the dwell time (D [sec]). To calculate the offline ITR, U in (3.7) was defined as follows:

$$U_{\mathrm{ETI}}^{\mathrm{off}} = I + \overline{T} + D, \tag{3.10}$$

where I [sec] is the instructing period and $\overline{T}$ is as follows:

$$\overline{T} = \frac{1}{N} \sum_{i=1}^{N} T_i, \tag{3.11}$$

where T_i [sec] is the duration from the start of the gazing period to the time when the gaze point lands on a target, and N is the number of trials for each size. We set $I = 1.0$ sec, $D = 0.5, 0.6, \ldots, 1.8, 1.9$ sec, and $N = 20$.

Additionally, to calculate the simulated online ITR, U in (3.7) was defined as follows:

$$U_{\mathrm{ETI}}^{\mathrm{on}} = G + \overline{Z}, \tag{3.12}$$

where G [sec] is the time for shifting the gaze from one target to another, and $\overline{Z}$ is as follows:

$$\overline{Z} = \frac{1}{N}\sum_{i=1}^{N} Z_i, \tag{3.13}$$

where Z_i $(\geq D)$ [sec] is the duration from the time when the gaze point first lands on the target to the time when a command is input, as illustrated in figure 3.6. The duration Z_i is dependent on the dwell time D and the subjects' performance. We set $G = 0.5$ sec.

Remarks on time window and dwell time

As described above, the shortest analysis time, the time window W, and the dwell time D were set as 0.5 sec; however, they could be shorter than 0.5 sec in the offline analysis. We did not consider such a short time here, since in practical interfaces, a very short interval for command inputs might cause an undesired repeat of the same command, such as a *press-and-hold* buttons or keys. Likewise, in the community of ETIs, it is called a *Midas touch problem*, which causes unintentional inputs using a very short dwell time; that is, every target a user looks at will be selected regardless of whether it is voluntary or not [81]. The shortest analysis time, 0.5 sec, was set following the default duration until a press turns into a long press in Android [82].

3.2.7 Statistical analysis

The statistical analyses, including an analysis of variance (ANOVA), a regression analysis, and a Wilcoxon signed-rank test, were conducted on the experimental results. A two-way between subjects ANOVA was performed

to compare the effect of the target size and the analysis time on the offline ITR for each BMI, ETI-target, and ETI-nearest. A multiple linear regression analysis was also performed to study how the independent variables (i.e., the target size and the analysis time) were related to the offline ITR. Lastly, paired-sample, one-sided Wilcoxon signed-rank tests were performed to compare the best performance of the BMI with that of the ETI-target and the ETI-nearest. The compared performance included the accuracy, the offline ITR, and the simulated online ITR.

3.3 Results

3.3.1 Experimental results

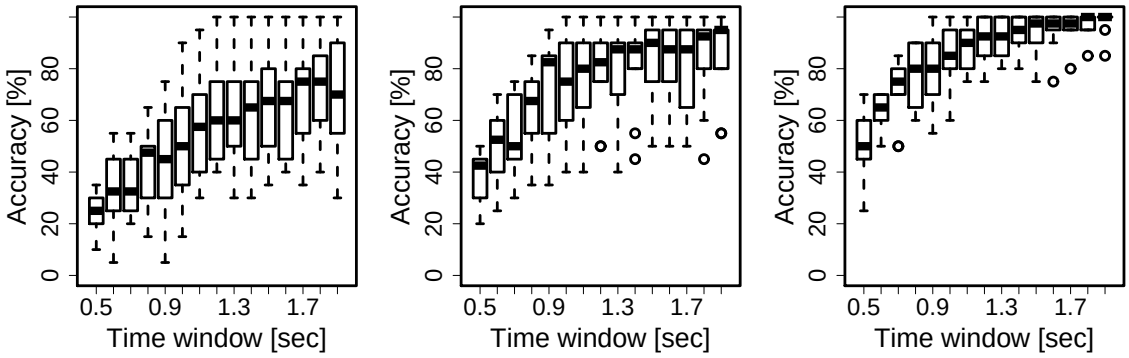
Recognition accuracy

Figures 3.7, 3.8, and 3.9 show the box plots of the distributions of the recognition accuracy of the BMI, the ETI-target, and the ETI-nearest, respectively. The horizontal axis shows the command analysis time (i.e., the time window W for the BMI and the dwell time D for the ETI), and the vertical axis shows the accuracy. Moreover, (a), (b), and (c) in the figures show the results when the target sizes were 20 mm, 40 mm, and 60 mm, respectively. The band inside the box represents the median.

Figures 3.7, 3.8, and 3.9 illustrate that, in terms of the target size, the larger target ($d = 60$ mm) showed higher accuracy than the smaller one ($d = 20$ mm) for both the BMI and the ETI. Another aspect is the analysis time (i.e., the time window for the BMI and the dwell time for the ETI). It can be seen from these figures that inappropriate analysis times yielded low accuracy. However, it is worth noting that the behavior of accuracy, as it relates to the analysis time, was completely different for each interface, as seen in figures 3.7, 3.8, and 3.9. The BMI showed higher accuracy when the time window

was longer, while the ETI-target showed higher accuracy with the shorter dwell time. The accuracy of ETI-nearest barely changed with respect to the dwell time.

In summary, the BMI had the highest mean accuracy with the long time window ($W = 1.9$ sec) and the largest target ($d = 60$ mm) (see figure 3.7c), and the ETI-target and the ETI-nearest had the highest mean accuracy with the shortest dwell time ($D = 0.5$ sec) and the largest target ($d = 60$ mm) (see figures 3.8c and 3.9c). The recognition accuracy with each target size is summarized in table 3.3. As shown in table 3.3, the mean and median accuracy of the BMI were higher than those of the ETI-target with all the target sizes, especially with the smallest target ($d = 20$ mm), in which the mean accuracy of the BMI achieved 72.0%, which is much higher than the highest mean accuracy of the ETI-target, 38.0%. Conversely, the ETI-nearest performed better than the BMI with the target size d = 20 and 40 mm, while the BMI achieved higher accuracy with the largest target ($d = 60$ mm). The statistical comparisons are described in section 3.3.2.

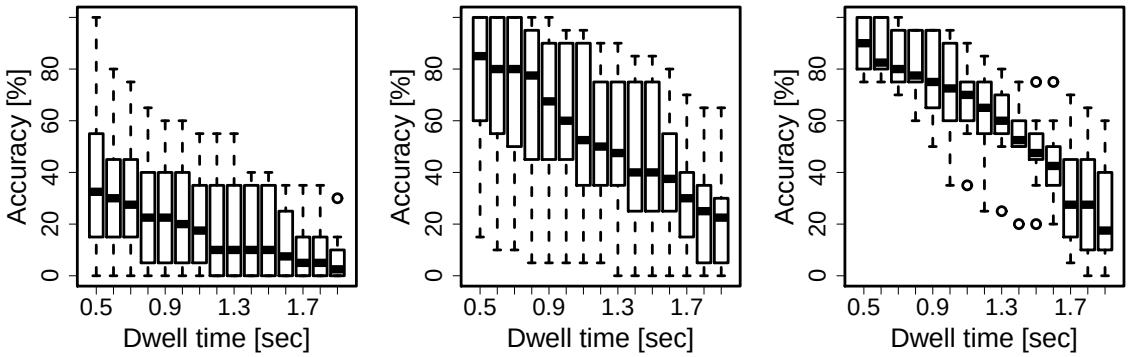


(A) Target size d = 20 mm

(B) Target size d = 40 mm

(C) Target size d = 60 mm

FIGURE 3.7: Box plots of the accuracy of the BMI.

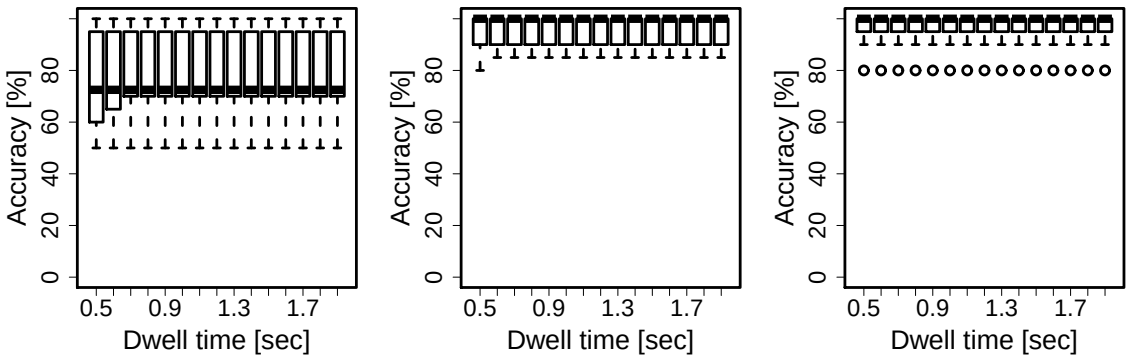


(A) Target size d = 20 mm

(B) Target size d = 40 mm

(C) Target size d = 60 mm

FIGURE 3.8: Box plots of the accuracy of the ETI-target.



(A) Target size d = 20 mm

(B) Target size d = 40 mm

(C) Target size d = 60 mm

FIGURE 3.9: Box plots of the accuracy of the ETI-nearest.

TABLE 3.3: Results of the recognition accuracy with each analysis time that led to the highest mean accuracy of the BMI, the ETI-target, and the ETI-nearest and the corresponding time window W and dwell time D. Significant differences are denoted by $*$ for $p < 0.05$.

Target size [mm]	Interface	W or D [sec]	Accuracy [%]		Significant difference
			Mean $\pm$ S.D.	Median	
d = 20	BMI	1.8	72.0 $\pm$ 19.5	75.0	$*$ ($>$ ETI-target)
	ETI-target	0.5	38.0 $\pm$ 30.1	32.5	
	ETI-nearest	0.7	77.5 $\pm$ 17.0	72.5	
d = 40	BMI	1.9	85.5 $\pm$ 17.2	95.0	
	ETI-target	0.5	74.5 $\pm$ 29.2	85.0	
	ETI-nearest	0.6	96.0 $\pm$ 5.7	100.0	$*$ ($>$ BMI)
d = 60	BMI	1.9	98.0 $\pm$ 4.8	100.0	$*$ ($>$ ETI-target)
	ETI-target	0.5	89.5 $\pm$ 9.6	90.0	
	ETI-nearest	0.5	96.5 $\pm$ 6.7	100.0	

ITR

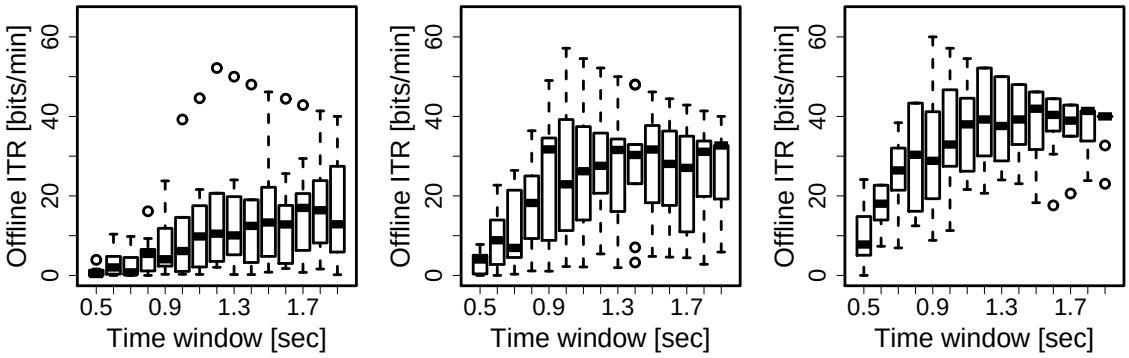
Figures 3.10, 3.11, and 3.12 show the box plots of the distributions of the offline ITR of the BMI, the ETI-target, and the ETI-nearest, respectively. The horizontal axis shows the command analysis time (i.e., the time window W for the BMI and the dwell time D for the ETI), and the vertical axis shows the offline ITR. Moreover, (a), (b), and (c) in figures 3.10, 3.11, and 3.12 show the results when the target sizes were 20, 40, and 60 mm, respectively.

Figures 3.10, 3.11, and 3.12 illustrate that, in terms of the target size, the larger target ($d = 60$ mm) led to higher ITR than the smaller one ($d = 20$ mm) for both the BMI and the ETI. These results were similar to the behavior of the recognition accuracy described above. The ETI-target and the ETI-nearest tended to have lower offline ITR when the dwell time was longer, as seen in figures 3.11 and 3.12. By contrast, it seemed that the BMI tended to have higher offline ITR when the time window was longer; however, it reached its peak at some point, as can be seen in figure 3.10.

In summary, the BMI had the highest mean offline ITR with the time window $W = 1.2$ sec and the largest target ($d = 60$ mm) (see figure 3.10c), and the ETI-target and the ETI-nearest had the highest mean offline ITR with the shortest dwell time ($D = 0.5$ sec) and the largest target ($d = 60$ mm) (see figures 3.11c and 3.12c). The offline ITRs with each target size are summarized in table 3.4. As shown in table 3.4, the mean and median offline ITR of the BMI were higher than those of the ETI-target using the smallest target ($d = 20$ mm). By contrast, the ETI-target showed better performance when using the medium and largest targets (d = 40 and 60 mm). The ETI-nearest achieved the highest offline ITR of the three regardless of the target size. The statistical comparisons are described in section 3.3.2.

For the analysis times that led to the highest mean ITR for the BMI, the ETI-target, or the ETI-nearest, the simulated online ITR was also estimated,

as described in section 3.2.6. The simulated online ITRs with each target size are summarized in table 3.5. By the definition of U in (3.7), the simulated online ITRs were generally higher than the offline ITRs. As in the offline ITR, the mean and median simulated online ITRs of the BMI were higher than those of the ETI-target using the smallest target ($d = 20$ mm). By contrast, the ETI-target showed better performance when using the medium and the largest targets (d = 40 and 60 mm), and the difference was much greater than that in the offline evaluation. Moreover, the simulated online ITR of the ETI-nearest was the highest of the three regardless of the target size. The statistical comparisons on the simulated online ITR are also described in section 3.3.2.

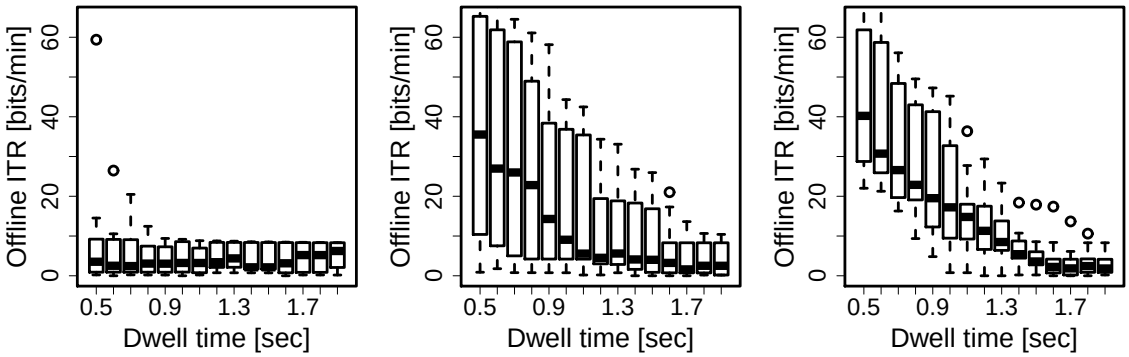


(A) Target size d = 20 mm

(B) Target size d = 40 mm

(C) Target size d = 60 mm

FIGURE 3.10: Box plots of the offline ITR of the BMI.

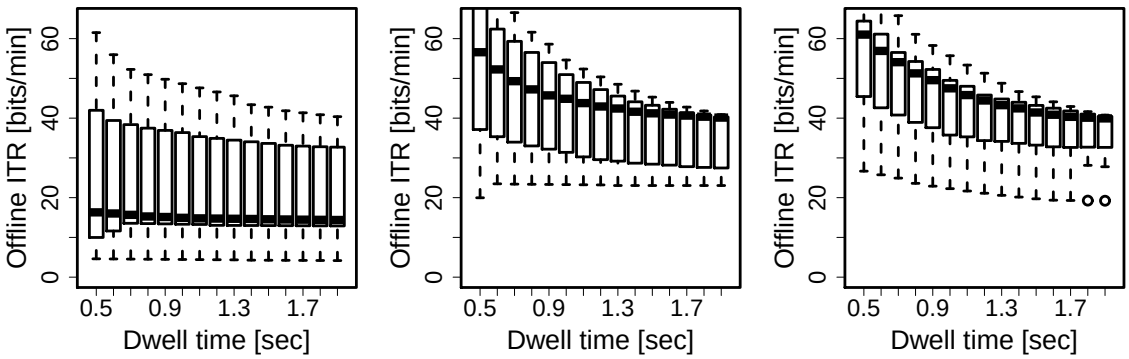


(A) Target size d = 20 mm

(B) Target size d = 40 mm

(C) Target size d = 60 mm

FIGURE 3.11: Box plots of the offline ITR of the ETI-target.



(A) Target size d = 20 mm

(B) Target size d = 40 mm

(C) Target size d = 60 mm

FIGURE 3.12: Box plots of the offline ITR of the ETI-nearest.

TABLE 3.4: Results of the offline ITR with each analysis time that led to the highest mean offline ITR of the BMI, the ETI-target, and the ETI-nearest and the corresponding time window W and dwell time D. Significant differences are denoted by $*$ for $p < 0.05$ and $**$ for $p < 0.01$.

Target size [mm]	Interface	W or D [sec]	Offline ITR [bits/min]		Significant difference
			Mean ± S.D.	Median	
d = 20	BMI	1.8	17.3 ± 12.4	16.4	$*$ (>ETI-target)
	ETI-target	0.5	10.0 ± 18.0	3.5	
	ETI-nearest	0.5	25.8 ± 20.5	16.3	
d = 40	BMI	1.2	28.8 ± 16.3	27.6	
	ETI-target	0.5	36.5 ± 29.3	35.5	
	ETI-nearest	0.5	53.8 ± 18.6	56.9	$**$ (> BMI)
d = 60	BMI	1.2	39.4 ± 10.9	39.2	
	ETI-target	0.5	44.9 ± 18.4	40.2	
	ETI-nearest	0.5	56.3 ± 14.5	61.0	$**$ (> BMI)

TABLE 3.5: Results of the simulated online ITR with each analysis time that led to the highest mean ITR in the offline evaluation. Significant differences are denoted by $*$ for $p < 0.05$ and $**$ for $p < 0.01$.

Target size [mm]	Interface	W or D [sec]	Simulated ITR [bits/min]		Significant difference
			Mean ± S.D.	Median	
d = 20	BMI	1.8	21.9 ± 15.6	20.7	$*$ (> ETI-target)
	ETI-target	0.5	15.9 ± 29.6	5.4	
	ETI-nearest	0.5	36.3 ± 32.4	20.5	
d = 40	BMI	1.2	38.9 ± 22.0	37.3	
	ETI-target	0.5	58.5 ± 46.4	55.8	
	ETI-nearest	0.5	81.9 ± 32.7	80.6	$**$ (> BMI)
d = 60	BMI	1.2	53.3 ± 14.7	53.1	
	ETI-target	0.5	72.5 ± 27.8	67.6	
	ETI-nearest	0.5	86.7 ± 22.1	87.9	$**$ (> BMI)

3.3.2 Statistical verifications

ANOVA and regression analysis

First, we describe the results of the ANOVA on the offline ITR of the BMI. There was a significant main effect of target size, $F(2, 405) = 124.38$, $p < 0.001$; also, there was a significant main effect of analysis time, that is, the time window length of EEG analysis, $F(14, 405) = 11.45$, $p < 0.001$. Meanwhile, there was no significant interaction of the target size and the analysis time, $F(28, 405) = 0.49$, $p = 0.987$. These results suggest that the target size and the analysis time do have an effect on the offline ITR of the BMI.

A multiple linear regression analysis was also performed to study how those independent variables (i.e., target size and analysis time) related to the offline ITR of the BMI. Here, we considered the model using target size d and analysis time (time window) W as explanatory variables given as:

$$\mathrm{ITR}_{\mathrm{BMI}} = \alpha + \beta_1 \cdot d + \beta_2 \cdot W. \tag{3.14}$$

Fitting the model above, we obtained:

$$\mathrm{ITR}_{\mathrm{BMI}} = -16.11 + 6.60d + 13.89W, \tag{3.15}$$

with an adjusted $R^2 = 0.435$. With this equation, it is presumed that the target size and the time window were positively related to the offline ITR of the BMI. Summarizing the results of the ANOVA and the regression analysis, our results suggest that the ITR of BMI is expected to be higher when using a large target or long time window.

Next, we describe the results of the ANOVA on the offline ITR of the ETI-target. There was a significant main effect of target size, $F(2, 405) = 43.34$, $p < 0.001$; in addition, there was a significant main effect of analysis time, that is, the dwell time, $F(14, 405) = 14.85$, $p < 0.001$. In addition, there

was a significant interaction between the target size and the analysis time, $F(28, 405) = 3.18$, $p < 0.001$. These results suggest that the target size, the analysis time, and the interaction of these factors do have an effect on the offline ITR of the ETI-target. The difference between the results of the ETI-target and the BMI was the significance of the interaction of target size and analysis time.

Again, a multiple linear regression analysis was performed to study how the target size and the analysis time were related to the offline ITR of the ETI-target. Here, we considered the model using target size d and analysis time D as explanatory variables given as:

$$\mathrm{ITR}_{\mathrm{ETI-target}} = \alpha + \beta_1 \cdot d + \beta_2 \cdot D. \tag{3.16}$$

Fitting the model above, we obtained:

$$\mathrm{ITR}_{\mathrm{ETI-target}} = 22.65 + 3.65d - 18.40D, \tag{3.17}$$

with an adjusted $R^2 = 0.327$. With this equation, it is presumed that the target size was positively related to the offline ITR of the ETI-target, while the dwell time was negatively related. To summarize the results of the ANOVA and the regression analysis, our results suggest that the ITR of ETI-target is expected to be higher when using a large target or short dwell time.

Lastly, we describe the results of the ANOVA on the offline ITR of the ETI-nearest. There was a significant main effect of target size, $F(2, 405) = 123.96$, $p < 0.001$; also, there was a significant main effect of analysis time, that is, the dwell time, $F(14, 405) = 4.23$, $p < 0.001$. In addition, there was no significant interaction between the target size and the analysis time, $F(28, 405) = 0.46$, $p = 0.993$. These results suggest that the target size and the analysis time do have an effect on the offline ITR of the ETI-nearest.

Again, a multiple linear regression analysis was also performed using target size d and analysis time D as explanatory variables as follows:

$$\mathrm{ITR}_{\mathrm{ETI-nearest}} = \alpha + \beta_1 \cdot d + \beta_2 \cdot D. \tag{3.18}$$

Fitting the model above, we obtained:

$$\mathrm{ITR}_{\mathrm{ETI-nearest}} = 27.35 + 6.14d - 10.05D, \tag{3.19}$$

with an adjusted $R^2 = 0.341$. With this equation, as in the ETI-target, it is presumed that target size was positively related to the offline ITR of the ETI-nearest, while dwell time was negatively related. To summarize the results of the ANOVA and the regression analysis, our results suggest that the ITR of ETI-nearest would be higher when using a large target or short dwell time.

Wilcoxon signed-rank test

The results of the Wicoxon tests are summarized in tables 3.3, 3.4, and 3.5. Only significant results reporting, the BMI showed significantly higher performance in the accuracy, the offline ITR, and the simulated online ITR compared to the ETI-target when using the smallest target ($d = 20$ mm). In addition, the accuracy of the BMI was significantly greater than that of the ETI-target when using the largest target ($d = 60$ mm). By contrast, when using the medium target ($d = 40$ mm), the ETI-nearest had significantly higher accuracy than the BMI. It also showed significantly higher offline ITR and simulated online ITR when using the medium and the largest targets (d = 40 and 60 mm).

3.4 Discussion

A comparison study on the similar design of BMI and eye tracking interfaces ETI was conducted in this chapter. Specifically, the SSVEP-based BMI and the dwelling-based ETI were compared. By investigating which interface is most appropriate under several conditions, the study ultimately aimed to improve the design of user interfaces. The performances of the SSVEP-BMI and the dwelling-ETI were evaluated by measuring the recognition accuracy and the ITR with respect to the target size and the command analysis time, that is, the time window length of EEG for the BMI and the dwell time for the ETI. The performance of the ETI was evaluated in two ways: 1) direct target selection (ETI-target) and 2) nearest target selection (ETI-nearest). Through the experiment of visual target selection tasks, we proved that the BMI was comparable to typical ETI (i.e., the ETI-target) in recognition accuracy and the ITR. Moreover, the ETI-nearest showed much better performance than the BMI, especially when the targets were large.

3.4.1 Tendencies toward the target size

To design a user interface using human sight, it is beneficial to appreciate behavior as it relates to target size because there is always limited space for displaying targets. In the experiment, three different target sizes were used (i.e., d = 20, 40, and 60 mm) for comparison.

For the BMI, the ETI-target, and the ETI-nearest, the accuracy and the offline ITR increased with a larger target. The decrease of the accuracy of the ETI-nearest when the target size was smaller may have been caused by the instruction to the subjects; we asked them to gaze at one of the targets gathered at the center of the display. The smaller the size of the targets, the closer to one another the targets were. For BMIs, a previous study [83, 84] also indicated that VEP response increased when the stimulus field (i.e., target size)

was enlarged. For typical ETIs, such as the ETI-target, other studies [35, 57, 85] have also indicated that the time to input each command and error rate decreases when the target size is enlarged. In addition, a previous study [86] claimed that accuracy generally increases as the target size expands, and the appropriate size of the target is dependent on the eye tracking device. Considering these previous studies, it is likely that the interfaces showed higher ITR with a larger target because higher accuracy results in higher ITR, according to equation (3.7).

3.4.2 Tendencies toward the analysis time

Another aspect of designing a user interface is setting the time to input a command, which is affected by the command analysis time. In most cases, a shorter analysis time is preferable, although an inappropriate analysis time leads to lower accuracy. In this study, the time window lengths of the EEG signal (i.e., W for the BMI) and the dwell time (i.e., D for the ETI) were set as $W = D = 0.5, 0.6, \ldots, 1.8, 1.9$ sec.

The BMI tended to have higher accuracy when the time window was longer. This result concurs with the results of previous studies [76, 72, 70]. It is worth noting that the accuracy would be saturated at some point in time, and the ITR would achieve the highest at that point and begin to decrease after according to equation (3.7). For this reason, the ITR is expected to increase as the time window gets longer; however, it will start to decrease after the time point when the accuracy is saturated. This tendency of the ITR has been indicated in previous studies [72, 70]. Contrary to the BMI, the ETI-target tended to show higher accuracy and ITR when the dwell time was shorter, while the accuracy of the ETI-nearest barely changed with respect to the dwell time. Previous works [57, 81] have also claimed that longer dwell times are less suitable. Meanwhile, another previous study [78] suggested

that the dwell time should be longer than 0.5 sec for usability, claiming that the subjects disapproved the 0.5 sec of dwell time as it was too short.

3.4.3 Comparison of the recognition accuracy and the ITR

The BMI showed the best performance with the long time window, while the ETI showed the best performance with the short dwell time, as described in section 5.3. In this section, the comparison of the performance with each analysis time that led to the best performance is discussed. It is worth noting that the eye tracking system and the algorithm for command selection employed in the present study would be described as research grade, and products supplied by reliable enterprises would perform better than the ETI employed in the present study. Figure 3.13 draws a comparison between the BMI and the ETI in terms of the accuracy, the offline ITR, and the simulated online ITR achieved with each analysis time.

Recognition accuracy

As shown in table 3.3, the Wilcoxon signed-rank test showed that the BMI had significantly better performance than the ETI-target in terms of the recognition accuracy when the target sizes were the smallest and the largest (d = 20 and 60 mm). It can also be seen in figure 3.13a that most of the subjects had better performance using the BMI than using the ETI-target when the target sizes were the smallest and the largest (d = 20 and 60 mm). Although there was no significant difference, the BMI also showed better performance than the ETI-target when the target size was medium ($d = 40$ mm). It is equally important to mention that one subject out of ten achieved 100% of accuracy using the ETI-target with the shortest dwell time ($D = 0.5$ sec), even when the target size was the smallest ($d = 20$ mm). Thus, there are exceptions, however, the experimental results should generally hold true. This

result quantitatively suggests, on the whole, that the BMI has an advantage in terms of the recognition accuracy compared to the ETI-target. It seems possible that the misclassification of the ETI-target was caused by the offset error in calibration as well as an inappropriate dwell time. Hence, increasing the frequency of calibration would increase the accuracy of the ETI-target at the same level as the BMI. However, the BMI is superior to the ETI-target in terms of the fact that it does not need calibration. Comparing accuracies between the BMI and the ETI-nearest, the ETI-nearest achieved significantly higher accuracy with the medium target ($d = 40$ mm). We can also see from figure 3.13d that most of the subjects performed better with the ETI-nearest, especially when using the medium and the largest targets (d = 40 and 60 mm). However, as mentioned earlier in section 3.2.5, it should be kept in mind that the ETI-nearest may not perform as well as it did in the present study, especially when implementing a large number of commands densely located on a display. When the targets are located densely, the ETI-nearest could lead to accidental false entries.

ITR

The important factor for evaluating interfaces is not only the recognition accuracy but also the speed. The ITR allows us to evaluate both the recognition accuracy and the speed; hence, the ITR is also an important criterion. When using the smallest target ($d = 20$ mm), the BMI had significantly higher ITR than the ETI-target in the offline and the simulated online evaluations, as described in section 3.3.2. It also appears from figures 3.13b and 3.13c that most of the subjects had better performance using the BMI than using the ETI-target when the target size was small ($d = 20$ mm). Again, one subject out of ten achieved outstandingly high offline and simulated online ITR at 59.4 bits/min and 98.0 bits/min, respectively, using the ETI-target with the shortest dwell time ($D = 0.5$ sec) and the smallest target ($d = 20$ mm). This

result suggests that the ETI-target generally needs sufficiently large targets, and the BMI has an advantage in target size over the ETI-target, even though the performance might be different depending on the person. Conversely, when using the medium and the largest targets (d = 40 and 60 mm), the ETI-target showed better performance than the BMI in terms of the offline and the simulated online ITR even though there was no significant difference. As depicted in figures 3.13b and 3.13c, with the medium and the largest targets (d = 40 and 60 mm), half the subjects or more showed higher ITR using the ETI-target than using the BMI, and their ITRs using the ETI-target were much higher than when using the BMI. This result indicates that the ETI-target potentially has better performance than the BMI when the target size is large enough. Furthermore, the ETI-nearest showed significantly higher offline and simulated online ITR than the BMI with the medium and the largest targets (d = 40 and 60 mm). Almost all the subjects achieved higher ITR using the ETI-nearest than the BMI, as seen in figures 3.13e and 3.13f. It is indicated that the ETI-nearest would be preferable when the number of the targets is small or the targets are located sparsely.

3.4.4 Summary

It was observed that the SSVEP-based BMI showed better performance than the dwelling-based ETI-target (i.e., typical ETI) when the target size was small. However, the experimental results showed that the ETI-target had potentially better performance than the BMI when the target size was large enough. In addition, it was indicated that the ETI-nearest would perform much better than the BMI when the targets are not located densely on the display. In the present study, the distance between each target was changed with the target size. Therefore, the experimental results are applicable when the screen size is small or large. These results suggest that the selection of

the SSVEP-based BMI or the dwelling-based ETI depends on the size of the screen that displays the targets. When a large screen, such as a monitor for a desktop computer, is available, the dwelling-based ETI is expected to have high accuracy and high ITR. Otherwise, when using small screens, such as smartphone or tablet screens, the SSVEP-based BMI would be more appropriate than the dwelling-based ETI.

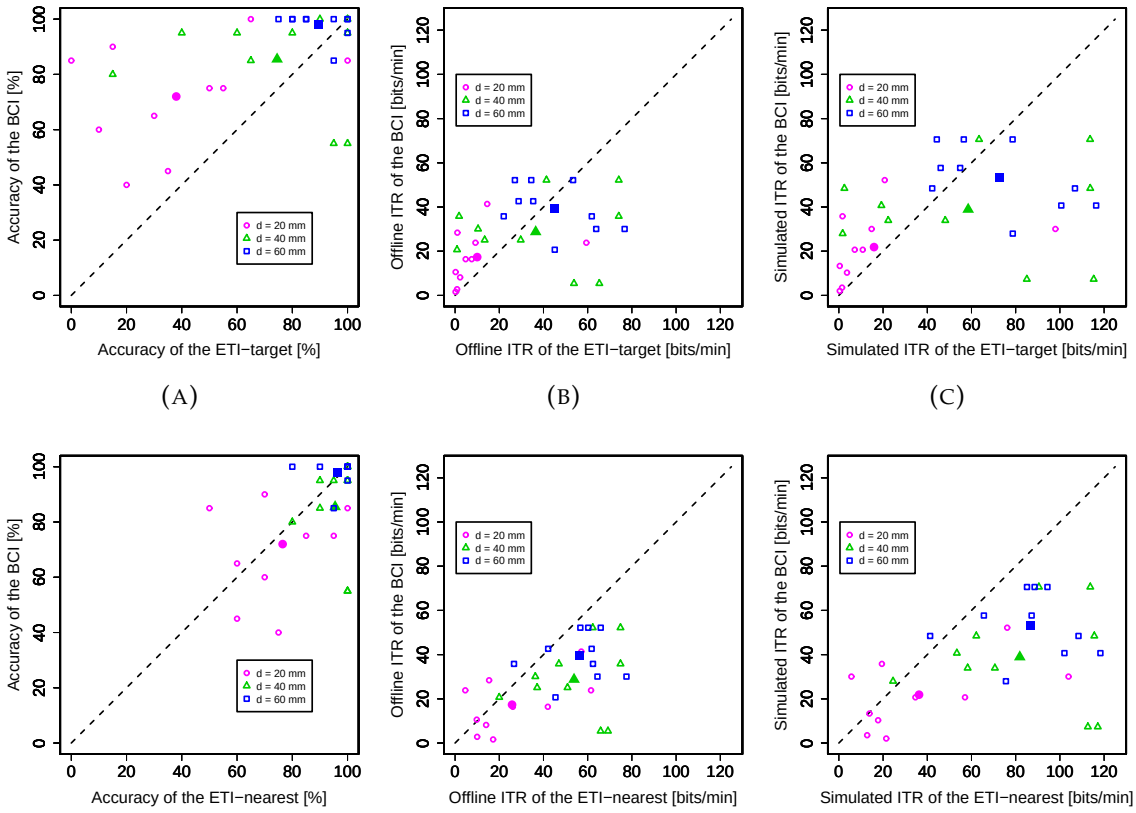


FIGURE 3.13: Comparisons of the recognition accuracy (13a and 13d), the offline ITR (13b and 13e), and the simulated online ITR (13c and 13f) with each analysis time that led to the highest mean value as listed in tables 3.3, 3.4, and 3.5. Figures 13a, 13b, and 13c represent comparisons between the BMI and the ETI-target, and figures 13d, 13e, and 13f represent comparisons between the BMI and the ETI-nearest. The values of each subject using the ETI-target or the ETI-nearest are plotted against those of the same subject using the BMI. The open circles, triangles, and squares represent the values with each target size, and the closed circles, triangles, and squares represent the mean values across the subjects. It is worth noting that the number of plots seem less than the number of subjects, as some subjects showed the same performance as the others and the plots overlapped.

Chapter 4

Calibration Reduction

Among SSVEP-BMIs, a state-of-the-art is the one based on hybrid frequency and phase coded SSVEP, which needs a large set of calibration data as reference signals, so-called individual templates. For hybrid frequency and phase coded SSVEP-BMIs, algorithms utilizing calibration data as reference signals, so-called individual templates, have recently been reported. The aim of this study is to propose an approach to calibration reduction by generating from individual templates corresponding to a part of commands (*source templates*) to new templates corresponding to the rest of commands. The new templates can be obtained by shifting the frequency and phase of the source template to the desired frequency and phase. In this way, time and cost for calibration can be greatly reduced. The experimental results suggested that the proposed approach successfully transferred the source template, closely achieving the performance using the full calibration dataset.

4.1 Introduction

Since SSVEPs are phase-locked periodic with the same or integer multiples of the frequency of visual stimulus, multiple commands can be implemented on SSVEP-BMIs by assigning different frequencies and/or phases to each stimulus. The method which assigns both frequencies and phases is called hybrid frequency and phase coding [87]. In terms of hybrid frequency and phase

coding, there are two types of strategies: *mixed frequency and phase coding* and *joint frequency and phase coding*. Mixed coding implements $N_f \times N_p$ visual stimuli in total by assigning N_f different frequencies with each phase and N_p different phases with each frequency to a visual stimulus [88]. Besides, joint coding implements N_f visual stimuli in total by assigning N_f different frequencies to each visual stimulus and assigning different phases to two adjacent stimuli [87, 89]. Both strategies have been gaining attention [37, 26, 90], and it has been reported that both can achieve a high performance with a large number of commands [87].

Simultaneously with increasing the number of visual stimuli presented by the hybrid frequency and phase coding, algorithms for recognizing the frequency and phase of SSVEPs have also been improved [37, 26, 90]. A previous study on a mixed-coded SSVEP-BMI [37] proposed a method utilizing a calibration data, which significantly outperformed other canonical correlation analysis (CCA)-based methods. In the method, an average of EEG signals in a calibration dataset across trials, so-called an individual template or a training reference signal, is used as a reference signal which efficiently characterizes temporal features of each individual's SSVEPs. Recently, further improved algorithms based on the individual templates have been reported [26, 90]. A major issue which still remains on these algorithms is that they require a long time for calibration to obtain an acceptable individual template. The minimum time for calibration can be estimated as *a few seconds for one trial* $\times$ *the number of trials for each command* $\times$ *the number of commands*. For instance, with the settings in [26], the minimum time for calibration is 8 minutes (1 second for a trial $\times$ 12 trials for each command $\times$ 40 commands in total) for calibration without any rest in between sessions. Such a calibration step can take a lot of time and energy and might cause a visual fatigue before an actual use. To deal with the problem, Yuan *et al.* proposed an approach for joint-coded SSVEP-BMIs exploiting inter-subject information [91]. The

method exploits the existing calibration data which were recorded beforehand by multiple users and averages their EEG signals to obtain templates for a new user, which enabled the new user to achieve a sufficient accuracy without any calibration. Although the method does not require any calibration for a new user, it is still necessary to have an access to an enormous calibration dataset of multiple users recorded in the same situation including electrode positions, visual stimuli, and so on.

To overcome this issue, we utterly propose a novel approach to calibration reduction by transferring the individual template of a certain command, called a *source template*, to obtain new templates of the other commands, called *target templates*. The target templates are generated by shifting the frequency and phase of the source template to the desired frequency and phase. An experiment with a mixed-coded SSVEP-BMI was conducted to evaluate performance of the proposed method in terms of the recognition accuracy.

4.2 Materials and Methods

4.2.1 Data Acquisition

Ten males and one female in their twenties took part in our experiment. All subjects were healthy and had normal or corrected-to-normal vision. They were given an informed consent, and this study was approved by the research ethics committee of Tokyo University of Agriculture and Technology.

A 23 inch LCD screen (BenQ, XL2411T) with a refresh rate of 120 Hz was used for displaying visual stimuli. During the experiment, the subjects seated on a comfortable chair in front of the screen about 60 cm away. In addition, the subjects ' heads were held steady using a chin rest. Figure 4.1 shows an image of displayed visual stimuli. As depicted in Figure 4.1, sixteen visual stimuli corresponding to each command were displayed on the

screen. All stimuli were approximately 5 centimeters on a side and disposed at equal intervals. Visual stimuli on each row flickered with frequencies of 12, 13, 14, or 15 Hz, respectively, and those on each column flickered with phases of 0, 90, 180, or 270 degrees, respectively. To present such flickering stimuli, an approximation approach *et al.*[92] was employed.

The subjects performed the following task in one trial. They started a trial by pressing the Enter key. One trial was configured with 1.5 seconds of pre-flickering interval, 4.0 seconds of flickering time, and 1.5 seconds of post-flickering interval. In pre-flickering interval, non-flickering visual stimuli and a white rectangle were displayed. In flickering time, all stimuli started flickering and the subjects were asked to gaze at the instructed stimulus by the white rectangle. In post-flickering interval all stimuli stopped flickering. After each trial, the task was stopped and the subjects were asked to rest their eyes. The task was restarted when the subjects pressed the Enter key again. 20 trials were performed with respect to each stimulus in a random order. Thus, each subject performed 320 trials in total.

To record EEG signals, we used Ag/AgCl active electrodes of Guger Technologies (g.tec) named g.LADYbird, g.LADYbirdGND (for GND), and g.GAMMAearclip (for reference, ear-clip type) for recording EEG signals. These were driven by the power supply unit named g.GAMMAbox (g.tec). Twenty five electrodes in accordance with the 10-5 system [93] were placed at CPz, CP1, CP2, CP3, CP4, CP5, CP6, Pz, P1, P2, P3, P4, P5, P6, P7, P8, POz, PO3, PO4, PO7, PO8, Oz, O1, O2 and Iz. The electrodes for GND and reference were positioned at AFz and A1, respectively.

4.2.2 SSVEP Recognition Using MCCA

The authors have recently proposed a novel method to recognize mixed-coded SSVEPs which achieved high performance [90]. The method employs

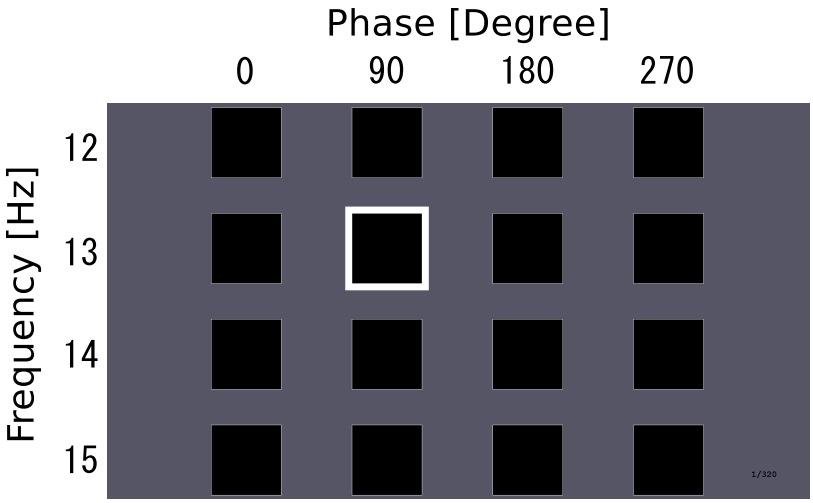


FIGURE 4.1: Image of presented visual stimuli that reversed black/white pattern. Subjects were instructed to gaze at a stimulus framed by a white rectangle.

multiset canonical correlation analysis (MCCA) [47] to obtain a spatial filter which enhances SSVEP components.

MCCA is a generalization of CCA to multiple datasets [47, 94, 95, 96, 97]. It finds weight vectors that maximize a correlation between weighted linear combinations of each dataset, called canonical variates. Let $\mathbf{S}_i \in \mathbb{R}^{I_i \times N}$ $(i = 1, \ldots, Q)$ be an I_i-channel signal that is normalized to have zero mean and unit variance. Their linear combinations, canonical variates, are denoted by $\mathbf{w}_i^T \mathbf{S}_i$, where $\mathbf{w}_i \in \mathbb{R}^{I_i \times 1}$ is a weight vector. With the MAXVAR criterion [47, 95, 96], MCCA solves the following maximization problem to find weight vectors $\mathbf{w} = [\mathbf{w}_1^T, \mathbf{w}_2^T, \cdots, \mathbf{w}_Q^T]^T$:

$$\rho = \max_{\mathbf{w}_1, \ldots \mathbf{w}_Q} \sum_{i \neq j}^{Q} \mathbf{w}_i^T \mathbf{S}_i \mathbf{S}_j^T \mathbf{w}_j \quad \text{s.t.} \ \frac{1}{Q} \sum_{i=1}^{Q} \mathbf{w}_i^T \mathbf{S}_i \mathbf{S}_i^T \mathbf{w}_i = 1. \tag{4.1}$$

Using a Lagrange multiplier technique, this objective function can be transformed into a generalized eigenvalue problem. The eigenvectors corresponding to the eigenvalues sorted in a descending order and the largest eigenvalue are denoted by $\mathbf{w}_i^{(1)}, \mathbf{w}_i^{(2)}, \ldots, \mathbf{w}_i^{(\Sigma_i I_i)}$ and λ, respectively. The largest eigenvalue can be interpreted as the similarity of those multiple datasets and

the eigenvectors corresponding to the large eigenvalues can be interpreted as the weight vectors that increase the correlation between these datasets.

In accordance with [90], MCCA was applied substituting the EEG signal of test data $\mathbf{x}$ for $\mathbf{S}_1$, the artificial reference signal $\mathbf{y}_k$ for $\mathbf{S}_2$, and the individual template $\mathbf{z}_k$ for $\mathbf{S}_3$, where the subscript, k, describes the index for stimuli. If the kth stimulus has a frequency of f_k, then the artificial signal $\mathbf{y}_k$ consisted of Fourier series of simulated stimulus signals given as

$$\mathbf{y}_k = [\sin(2\pi f_k t), \cos(2\pi f_k t), \ldots, \sin(2L\pi f_k t), \cos(2L\pi f_k t)]^T , \tag{4.2}$$

where L is the number of harmonics in the Fourier series, the first two components are the sinusoids of the fundamental frequency f_k and the others are harmonics. The number of harmonics was set to $L = 3$ in the present study. Additionally, the individual template $\mathbf{z}_k$ corresponding to the kth stimulus was obtained by averaging EEG signals in a calibration dataset across trials. Applying MCCA to these datasets, the largest eigenvalue λ_k and a spatial filter $\mathbf{w_x}$ that was an eigenvector corresponding to the largest eigenvalue were obtained. With the spatial filter, the test signal $\mathbf{x}$ and the individual template $\mathbf{z}_k$ were projected. Subsequently, correlation coefficients between these two projections were obtained as:

$$\rho_k = \mathrm{Corr}\left(\mathbf{w_x}^T\mathbf{x}, \mathbf{w_x}^T\mathbf{z}_k\right), \tag{4.3}$$

where $\mathrm{Corr}(\cdot,\cdot)$ is the Pearson's correlation coefficient. The correlation coefficients were computed for each command k $(k = 1, \ldots, K)$ following Equation (4.3). The largest eigenvalue and the correlation coefficient were combined and used for detecting a command as follows:

$$k^\star = \arg\max_k \ \lambda_k \mathrm{sign}\left(\rho_k\right) \cdot \left(\rho_k\right)^2 . \tag{4.4}$$

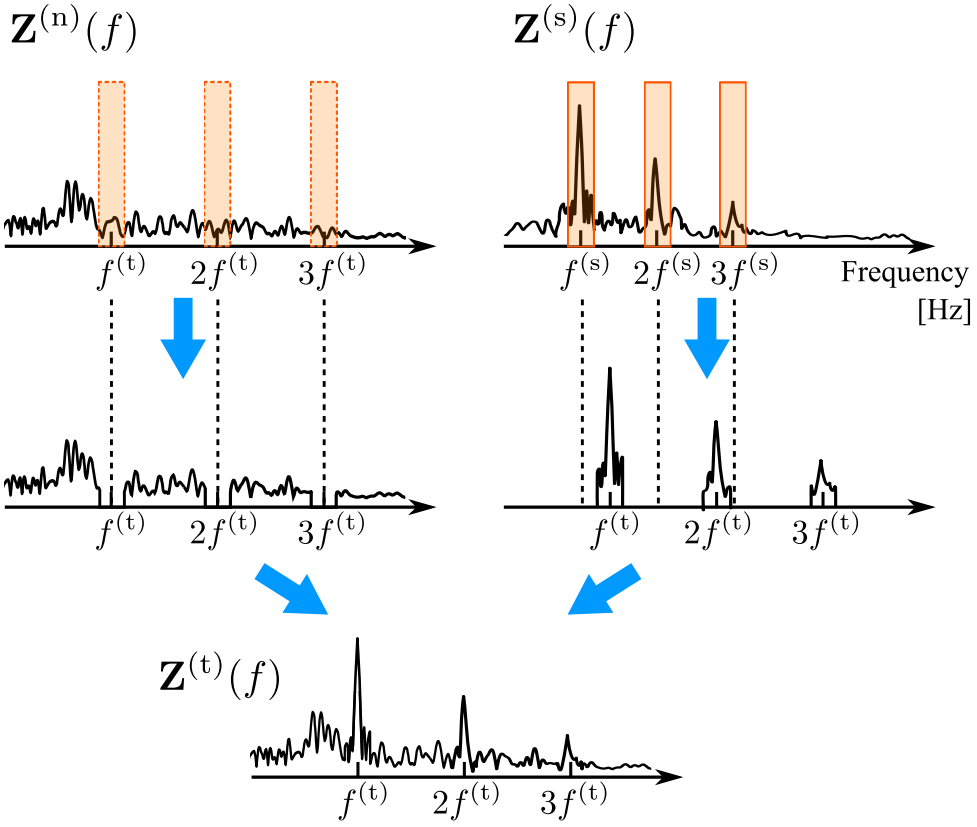


FIGURE 4.2: Frequency shifting by a bank of bandpass filters and a bank of bandstop filters. Fourier transforms of the non-control EEG signal, the source template, and the target template are denoted by $\mathbf{Z}^{(\mathrm{n})}(f)$, $\mathbf{Z}^{(\mathrm{s})}(f)$, and $\mathbf{Z}^{(\mathrm{t})}(f)$, respectively.

4.2.3 Transferring the Source Template

As described earlier, the individual templates can be obtained by averaging EEG signals across trials for each command. Necessarily, it requires a long time for calibration. To overcome this issue, we aimed to transfer the individual template of a certain command, named a source template, to obtain a new template of the other command, named a target template. The source template can be transferred to the target template by these two following approaches:

Shifting the phase A source template is shifted in time domain with a time delay which corresponds to a phase desired to be shifted;

Shifting the frequency A source template is Fourier transformed into a frequency domain. Subsequently, the peak at a source frequency is shifted to a target frequency and inverse transformed into a time domain.

Details for these approaches are described below.

Shifting the phase

Let $\mathbf{z}^{(\mathrm{s})}(t)$ be a source template given as an average of the recorded EEG signals as a response to a visual stimulus flickering with a certain frequency $f^{(\mathrm{s})}$ and a phase $\phi^{(\mathrm{s})}$. A target template $\mathbf{z}^{(\mathrm{t})}(t)$ with the same frequency $f^{(\mathrm{s})}$ as the source template and a target phase $\phi^{(\mathrm{t})}$ is obtained by shifting the source template with a time delay as follows:

$$\mathbf{z}^{(\mathrm{t})}(t) = \mathbf{z}^{(\mathrm{s})}(t + \tau), \tag{4.5}$$

where the time delay τ is defined as $\tau = \frac{\phi^{(\mathrm{t})} - \phi^{(\mathrm{s})}}{2\pi f^{(\mathrm{s})}}$.

Shifting the frequency

The following procedure is based on an assumption that base lines of Fourier transforms of the source template and the target template share a common characteristic, and that common characteristic would appear in EEG signals in non-control state. The EEG signals in non-control state, i.e., when the user is not gazing at any visual stimulus, would represent a background activity, a spontaneous activity, an environmental noise, and so forth. Accordingly, the target template in frequency domain can be ideally obtained by exploiting non-control EEG signals as a base line and shifting the peaks of the source template at a source frequency to a target frequency. Figure 4.2 illustrates an image of the proposed architecture of frequency shifting.

Let $\mathbf{Z}^{(\mathrm{s})}(f)$ and $\mathbf{Z}^{(\mathrm{n})}(f)$ be Fourier transforms of the source template and the non-control EEG signal, respectively. Firstly, Fourier transform of the

source template $\mathbf{Z}^{(\mathrm{s})}(f)$ was filtered by a bank of bandpass filters that passes the harmonics corresponding to the source frequency, $f^{(\mathrm{s})}$, of the source template. Specifically, the outputs of the filter bank are denoted by

$$\mathbf{Z}_1^{(\mathrm{s})}(f) = \mathcal{H}_{f^{(\mathrm{s})}}(f)\mathbf{Z}^{(\mathrm{s})}(f), \tag{4.6}$$

$$\vdots$$

$$\mathbf{Z}_L^{(\mathrm{s})}(f) = \mathcal{H}_{Lf^{(\mathrm{s})}}(f)\mathbf{Z}^{(\mathrm{s})}(f), \tag{4.7}$$

where L is the number of harmonics and $\mathcal{H}_{\overline{f}}(f)$ is the transfer function of a bandpass filter at the center frequency of $\overline{f}$ with a bandwidth of $2\Delta f$, defined as:

$$\mathcal{H}_{\overline{f}}(f) = \begin{cases} 1, & \overline{f} - \Delta f \leq f \leq \overline{f} + \Delta f \\ 0, & \text{otherwise} \end{cases}, \tag{4.8}$$

where Δf was set to 0.5 Hz in the present study. Secondly, Fourier transform of the non-control signal $\mathbf{Z}^{(\mathrm{n})}(f)$ was filtered by a bank of bandstop filters so as to exclude bands corresponding to the target frequency $f^{(\mathrm{t})}$ and its harmonics, and then, previously extracted peaks corresponding to the source frequency was shifted to a target frequency and added as follows:

$$\begin{aligned} \mathbf{Z}^{(\mathrm{t})}(f) = &\left\{\overline{\mathcal{H}}_{f^{(\mathrm{t})}}(f) + \overline{\mathcal{H}}_{2f^{(\mathrm{t})}}(f) + \cdots + \overline{\mathcal{H}}_{Lf^{(\mathrm{t})}}(f)\right\}\mathbf{Z}^{(\mathrm{n})}(f) \\ &+ \mathbf{Z}_1^{(\mathrm{s})}\left(f - \left(f^{(\mathrm{t})} - f^{(\mathrm{s})}\right)\right) \\ &+ \mathbf{Z}_2^{(\mathrm{s})}\left(f - \left(2f^{(\mathrm{t})} - 2f^{(\mathrm{s})}\right)\right) \\ &+ \cdots + \mathbf{Z}_L^{(\mathrm{s})}\left(f - \left(Lf^{(\mathrm{t})} - Lf^{(\mathrm{s})}\right)\right), \end{aligned} \tag{4.9}$$

where $\overline{\mathcal{H}}_{\overline{f}}(f)$ is the transfer function of a bandstop filter at the center frequency of $\overline{f}$ with a bandwidth of $2\Delta f$, defined as:

$$\overline{\mathcal{H}}_{\overline{f}}(f) = \begin{cases} 0, & \overline{f} - \Delta f \leq f \leq \overline{f} + \Delta f \\ 1, & \text{otherwise} \end{cases}. \tag{4.10}$$

Finally, the target template $\mathbf{z}^{(\mathrm{t})}(t)$ with the target frequency $f^{(\mathrm{t})}$ and the same phase $\phi^{(\mathrm{s})}$ as the source template was obtained by the inverse Fourier transform of $\mathbf{Z}^{(\mathrm{t})}(f)$.

Case study for transferring the source template

Consider a set of mixed-coded stimuli consisting of $N_f \times N_p$ commands in total, where N_f and N_p are the number of different frequencies and phases that are assigned to each visual stimulus. We assumed the following three cases:

Phase-shifted template (PST) As source templates, use EEG signals with respect to visual stimuli of N_f different frequencies. PST is obtained by shifting the phase of each source template.

Frequency-shifted template (FST) As source templates, use EEG signals with respect to visual stimuli of N_p different phases and one phase. PST is obtained by shifting the frequency of each source template.

Phase-and-frequency-shifted template (PFST) As a source template, use EEG signals with respect to a visual stimulus of one phase and one phase. PFST is obtained by shifting both the phase and frequency of the source template.

In this chapter, the individual templates with the frequency $f^{(\mathrm{s})} = 15$ Hz and $\phi^{(\mathrm{s})} = 180$ degrees were employed as the source templates to obtain the PSTs and FSTs, respectively. For the PFSTs, the individual template with the

TABLE 4.1: Averaged mixed frequency and phase recognition accuracy (MACC), frequency recognition accuracy (FACC), phase recognition accuracy (PACC) across the subjects using the data length of 2.0 seconds. The top three rows show the results based on the proposed approach, and the lower two row shows that based on the prior works.

Method	MACC [%]	FACC [%]	PACC [%]
MCCA w/ PST	85.17 ± 15.04	90.23 ± 11.39	86.68 ± 13.33
MCCA w/ FST	80.45 ± 16.72	89.69 ± 11.04	83.13 ± 13.98
MCCA w/ PFST	79.86 ± 18.38	89.12 ± 14.21	82.47 ± 15.00
MCCA w/ IT	91.31 ± 11.67	94.32 ± 8.97	92.39 ± 10.02
CCA	23.04 ± 2.25	92.81 ± 9.62	25.00 ± 0.00

frequency $f^{(\mathrm{s})} = 15$ Hz and $\phi^{(\mathrm{s})} = 180$ degrees was employed as the source template. The non-control signal was obtained by averaging EEG signals of pre-flickering interval. Furthermore, the phase was first shifted and the frequency was shifted next to obtain the PFSTs.

4.3 Results

To evaluate performance, mixed frequency and phase, frequency, and phase recognition accuracy were measured, respectively. These accuracies using MCCA with PST, FST, and PFST were compared with the previous works including MCCA with individual template (IT) [90] and the standard CCA method [39, 38] for the reference methods. It should be noted that MCCA with IT need a full calibration, while the standard CCA cannot decode a phase. The performance of these methods were evaluated based on 4-fold cross-validation.

Figures 4.3a, 4.3b, and 4.3c show the averaged recognition accuracies of mixed frequency and phase, frequency only, and phase only. Table 4.1 summarizes the averaged accuracies across the subjects using the data length of 2.0 seconds. As shown in Figures 4.3a, 4.3b, and 4.3c, MCCA with IT achieved the highest accuracy compared to the other methods in either case

of recognizing the mixed frequency and phase, the frequency, and the phase as expected. Among the proposed approaches, PST achieved a slightly higher accuracy compared to FST. Remarkably, PFST which required only one source template also performed well closely achieving to FST and PST as shown in Table 4.1. Besides, in terms of the frequency recognition, MCCA with PST, FST, and PFST was inferior to the standard CCA as presented in Figure 4.3b. However, it is worth noting that MCCA with PST, FST, and PFST can be applied to the phase recognition and the mixed frequency and phase recognition as well.

4.4 Discussions

The experimental results showed that the proposed transferred template enabled to recognize both frequency and phase with a small amount of calibration closely achieving the performance using the full calibration data. The difference between the highest accuracies using the transferred templates and the individual templates was -11.45%, while the minimum time for calibration could be decreased by 93.75%.

Although mixed-coded SSVEP-BMI was employed to evaluate the proposed transferring approach in this chapter, PFST can be applied to joint-coded SSVEP-BMIs likewise. This means that this study is potentially applicable to many of existing calibration-based algorithms [26, 37, 98] for recognition of the frequency and phase of SSVEPs. In the case of applying PFST, the minimum time for calibration would be reduced by $1/$*the number of commands in total*, which is an important improvement for a practical use.

As described in Section 6.3, the frequency recognition accuracy using the proposed approach was slightly lower than that of the standard CCA method [39]. Hence, further studies might explore how to shift the frequency in a

more sophisticated manner in order to achieve higher performance. Moreover, an online adaptation of the transferred templates [91] would also improve the performance and help us to establish a more reliable interface.

In conclusion, we proposed an innovative approach to calibration reduction for hybrid frequency and phase coded SSVEP-BMIs. The present results suggested the feasibility of the proposed approach to achieve a high performance with a small calibration.

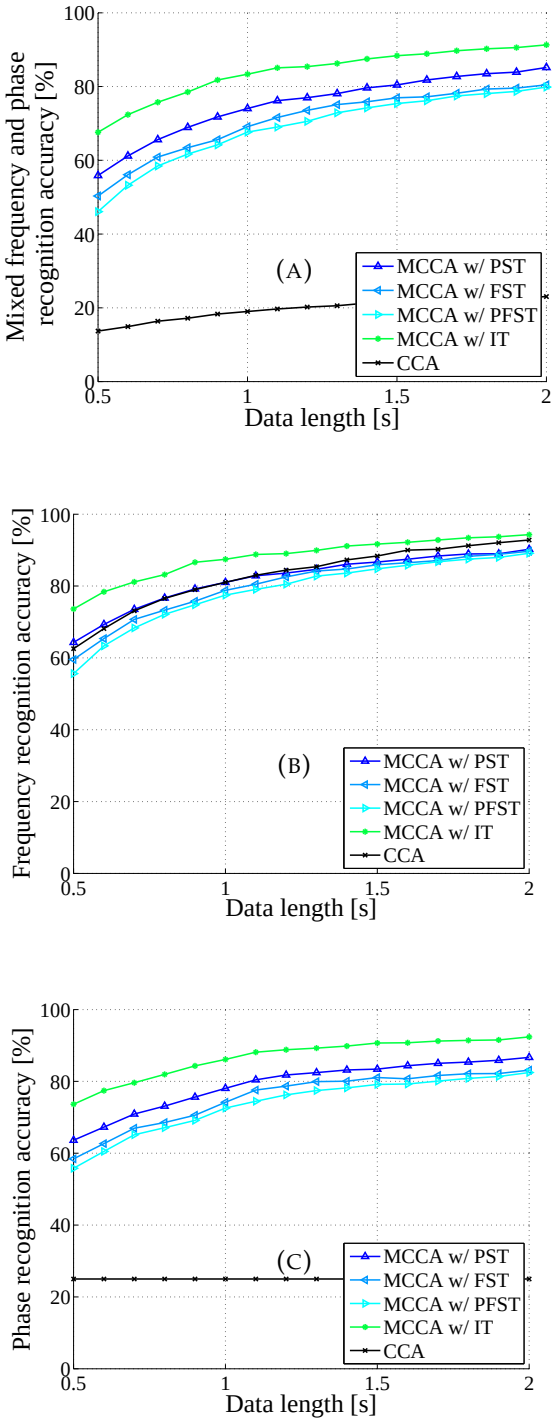


FIGURE 4.3: Averaged recognition accuracies across the subjects of (a) mixed frequency and phase, (b) frequency, and (c) phase. Blue lines with triangle markers shows the results based on the proposed approach, and the other two lines shows the results based on the prior works.

Chapter 5

Decoding Hybrid-Coded SSVEP

For SSVEP-based BMI, mixed frequency and phase coding which can implement a number of commands and achieve a high information transfer rate (ITR) has recently been gaining much attention. In order to implement mixed-coded SSVEP-BMI as a reliable interface, it is important to detect commands fast and accurately. This chapter presents a novel method to recognize mixed-coded SSVEPs which achieves high performance. The method employs multiset canonical correlation analysis to obtain spatial filters which enhance SSVEP components. An experiment with a mixed-coded SSVEP-BMI was conducted to evaluate performance of the proposed method compared with the previous work. The experimental results showed that the proposed method achieved significantly higher command recognition accuracy and ITR than the state-of-the-art.

5.1 Introduction

SSVEPs are phase-locked periodic with the same or integer multiples of the frequency of visual stimulus. Thus, frequency coding assigning different frequencies to each stimulus, phase coding assigning different phases to each stimulus, and mixed frequency and phase coding can be implemented on SSVEP-based BMIs [88].

Recently, a method of detecting mixed-coded SSVEPs has been proposed [37]. The method employed canonical correlation analysis (CCA)[38, 39] utilizing a training set, and enabled increasing a number of commands, improving a recognition accuracy, and boosting an information transfer rate (ITR) [2]. However, this CCA-based method requires to solve a generalized eigenvalue problem several times to obtain multiple spatial filters which enhance SSVEP components.

In this chapter, we propose to employ multiset canonical correlation analysis (MCCA) [47] for detecting mixed coded SSVEPs in order to simplify the CCA-based method [37]. Since MCCA enables us to evaluate a relationship between multiple datasets simultaneously, the proposed method requires to solve the generalized eigenvalue problem just once in order to obtain spatial filters. An experiment was conducted to evaluate performance of the proposed method in terms of the recognition accuracy and the ITR.

5.2 Materials and Methods

5.2.1 Data Acquisition

Ten males and one female in their twenties took part in our experiment. All subjects were healthy and had normal or corrected-to-normal vision. They were given an informed consent, and this study was approved by the research ethics committee of Tokyo University of Agriculture and Technology.

A 23 inch LCD screen (BenQ, XL2411T) with a resolution of 1,920 $\times$ 1,080 and a refresh rate of 120 Hz was used for displaying visual stimuli. During the experiment, the subjects seated on a comfortable chair in front of the screen about 60 cm away.

We used Ag/AgCl active electrodes of Guger Technologies (g.tec) named

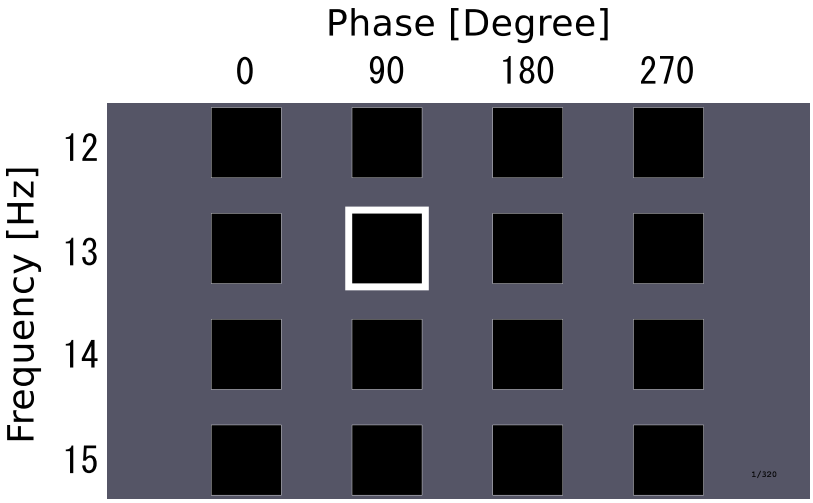


FIGURE 5.1: Image of displayed targets that reversed black/white pattern. Subjects were instructed to gaze at a target framed by a white rectangle.

g.LADYbird, g.LADYbirdGND (for GND), and g.GAMMAearclip (for reference, ear-clip type) for recording EEG signals. These were driven by the power supply unit named g.GAMMAbox (g.tec). Twenty five electrodes following the 10-5 system [93] were positioned at CPz, CP1, CP2, CP3, CP4, CP5, CP6, Pz, P1, P2, P3, P4, P5, P6, P7, P8, POz, PO3, PO4, PO7, PO8, Oz, O1, O2 and Iz. The electrodes for GND and reference were AFz and A1, respectively. The signals were amplified with a configuration of 50 μV/V (20,000 times) by a biosignal amplifier, MEG-6116 (Nihon Kohden) that had lowpass and highpass analog filter for each channel. Cut off frequencies of the low-pass and the high-pass filters were set to 100 Hz and 0.5 Hz, respectively. The amplified signal was sampled by an A/D converter, AIO-163202FX-USB (Contec), with a sampling rate of $1,200$ Hz. The sampled signals were recorded with Data Acquisition Toolbox of the MATLAB (MathWorks).

5.2.2 Design of Visual Stimuli

Figure 6.1 shows an image of displayed targets. As illustrated in Figure 6.1, sixteen visual stimuli corresponding to each command were displayed on the screen. All stimuli were about 5 centimeters square large (visual angle of 4.77 deg) and arranged at equal intervals. Visual stimuli on each row flickered with frequencies of 12, 13, 14, or 15 Hz, respectively, and visual stimuli on each column flickered with phases of 0, 90, 180, or 270 degree, respectively.

To produce such flickering stimuli described above, an approximation approach proposed by Wang *et al.*[92] was used. In this method, a sequence of visual flicker $s(f, \phi, i)$ at i-th frame with frequency f and phase ϕ as follows:

$$s(f, \phi, i) = \mathrm{square}[2\pi f(i/\mathrm{RefreshRate}) + \phi] \,, \tag{5.1}$$

where $\mathrm{square}[\cdot]$ denotes the operator that generates a 50% duty cycle square wave with levels 0 and 1.

5.2.3 Task

A task of of the subjects in a trial is described below. The subjects were asked to press the Enter key to start a trial. One trial consisted of 1.5 sec of pre-interval, 4.0 sec of flickering time, and 1.5 sec of post-interval. In pre-interval, not-flickering targets and a white rectangle were displayed. In flickering time, all stimuli started flickering and the subjects gazed at the instructed target by the white rectangle. In post-interval all stimuli stopped flickering. After each trial, the task was stopped and the subjects were asked to rest their eyes. The task was restarted when the subjects pressed the Enter key again. All target stimuli were indicated 20 times each in a random order. Hence, each subject performed 320 trials.

5.2.4 Data analysis for SSVEP recognition

EEG signals recorded at eight following channels, PO3, PO4, PO7, PO8, POz, O1, O2, and Oz that corresponded to parieto-occipital region were analyzed offline. The signals for 0.15 sec just after flickering started were neglected considering a delay of SSVEP onset [99]. Following two methods were applied to the signals to recognize SSVEPs and detect commands: the previous work based on CCA as a contrast method and the proposed method were applied to the signals.

CCA-based method [37]

Recently, a method for detecting SSVEP, which was based on CCA and utilized training data as a reference signal, has been proposed [37]. CCA is a statistical method to measure a relationship between two sets of multidimensional variables [38]. Let $\mathbf{S}_1 \in \mathbb{R}^{I_1 \times N}$ and $\mathbf{S}_2 \in \mathbb{R}^{I_2 \times N}$ be two multidimensional signals that are normalized to have zero mean and unit variance. Their weighted linear combinations, called canonical variates, are denoted by $\mathbf{w}_1^T \mathbf{S}_1$ and $\mathbf{w}_2^T \mathbf{S}_2$. CCA computes the weight vectors $\mathbf{w}_1$ and $\mathbf{w}_2$ that maximize the following maximization problem:

$$\rho = \max_{\mathbf{w}_1, \mathbf{w}_2} \frac{\mathbf{w}_1^T \mathbf{S}_1 \mathbf{S}_2^T \mathbf{w}_2}{\sqrt{\mathbf{w}_1^T \mathbf{S}_1 \mathbf{S}_1^T \mathbf{w}_1} \sqrt{\mathbf{w}_2^T \mathbf{S}_2 \mathbf{S}_2^T \mathbf{w}_2}}, \tag{5.2}$$

where ρ is called a canonical correlation coefficient. The optimal weight vectors $\mathbf{w}_1$ and $\mathbf{w}_2$ can be obtained by solving a generalized eigenvalue problem.

The method first finds three spatial filters which enhance the SSVEP components. The three spatial filters are obtained as weight vectors by applying CCA to all two-combinations of three multichannel signals which are a test EEG signal, an artificial reference signal, and a training reference signal, which are respectively denoted by $\mathbf{X} \in \mathbb{R}^{C \times N}$, $\mathbf{Y}_k \in \mathbb{R}^{2H \times N}$, and $\mathbf{Z}_k \in \mathbb{R}^{C \times N}$, where the subscript, k, describes the index for stimuli. If the kth stimulus

has a frequency of f_k, then the artificial signal $\mathbf{Y}_k$ consists of Fourier series of simulated stimulus signals given as

$$\mathbf{Y}_k = [\sin(2\pi f_k t), \cos(2\pi f_k t), \ldots, \sin(2H\pi f_k t), \cos(2H\pi f_k t)]^T \ , \tag{5.3}$$

where H is the number of harmonics in the Fourier series, the first two components are the sinusoids of the fundamental frequency f_k and the others are harmonics. Additionally, the training reference signal $\mathbf{Z}_k$ corresponding to the kth stimulus is obtained by averaging EEG signals of a training set across trials. In this chapter, $\mathbf{Z}_k$ was obtained based on leave-one-out cross-validation. Letting $\mathbf{S}_1$ be $\mathbf{X}$ and $\mathbf{S}_2$ be $\mathbf{Y}_k$ in (5.2), we obtain a spatial filter, $\mathbf{w}_1$, denoted by $\mathbf{w}_{\mathbf{XY}}$. In the same way, we can obtain $\mathbf{w}_{\mathbf{XZ}}$ from $\mathbf{X}$ and $\mathbf{Z}_k$ and $\mathbf{w}_{\mathbf{ZY}}$ from $\mathbf{Z}_k$ and $\mathbf{Y}_k$. Using these spatial filters, the correlation coefficients between the canonical variate of the test data $\mathbf{X}$ and that of the training reference signal $\mathbf{Z}_k$ are computed as follows:

$$\boldsymbol{\rho}_k = \begin{bmatrix} \rho_{k,1} \\ \rho_{k,2} \\ \rho_{k,3} \\ \rho_{k,4} \end{bmatrix} = \begin{bmatrix} \rho_k \\ \mathrm{Corr}\left(\mathbf{w}_{\mathbf{XY}}^T \mathbf{X}, \mathbf{w}_{\mathbf{XY}}^T \mathbf{Z}_k\right) \\ \mathrm{Corr}\left(\mathbf{w}_{\mathbf{XZ}}^T \mathbf{X}, \mathbf{w}_{\mathbf{XZ}}^T \mathbf{Z}_k\right) \\ \mathrm{Corr}\left(\mathbf{w}_{\mathbf{ZY}}^T \mathbf{X}, \mathbf{w}_{\mathbf{ZY}}^T \mathbf{Z}_k\right) \end{bmatrix} \ , \tag{5.4}$$

where ρ_k is the canonical correlation coefficient between $\mathbf{X}$ and $\mathbf{Y}_k$, and $\mathrm{Corr}(\cdot, \cdot)$ is the Pearson's correlation coefficient. The correlation coefficients are combined and used for detecting a command as follows:

$$k^\star = \arg\max_k \sum_{i=1}^{4} \mathrm{sign}(\rho_{k,i}) \cdot \rho_{k,i}^2 \ . \tag{5.5}$$

5.2.5 Proposed method based on multiset CCA

In CCA-based method [37], the generalize eigenvalue problem has to be solved several times to obtain the multiple spatial filters as described above. Therefore we propose to employ MCCA to obtain spatial filters that are given by solving the generalized eigenvalue problem just once.

MCCA is a generalization of CCA to multiple datasets [47]. It finds weight vectors that maximize a correlation between weighted linear combinations of each dataset, called canonical variates. To maximize the overall correlation, Kettenring [47] listed five possible objective functions. Here, we adopted the objective function called the MAXVAR since it is a natural extension of CCA to multiple datasets and can lead to a solution directly [96].

Let $\mathbf{S}_i \in \mathbb{R}^{I_i \times N}$ $(i = 1, \ldots, Q)$ be an I_i-channel signal that is normalized to have zero mean and unit variance. Their linear combinations, canonical variates, are denoted by $\mathbf{w}_i^T \mathbf{S}_i$, where $\mathbf{w}_i \in \mathbb{R}^{I_i \times 1}$ is a weight vector. With the MAXVAR criterion, MCCA solves following maximization problem to find weight vectors $\mathbf{w} = [\mathbf{w}_1^T, \mathbf{w}_2^T, \cdots, \mathbf{w}_Q^T]^T$:

$$\rho = \max_{\mathbf{w}_1, \ldots \mathbf{w}_Q} \sum_{i \neq j}^{Q} \mathbf{w}_i^T \mathbf{S}_i \mathbf{S}_j^T \mathbf{w}_j \quad \text{s.t.} \ \frac{1}{Q} \sum_{i=1}^{Q} \mathbf{w}_i^T \mathbf{S}_i \mathbf{S}_i^T \mathbf{w}_i = 1 \ . \tag{5.6}$$

Using a Lagrange multiplier technique, this objective function can be transformed into a generalized eigenvalue problem:

$$(\mathbf{R} - \mathbf{P})\mathbf{w} = \rho \mathbf{P} \mathbf{w} \ , \tag{5.7}$$

where

$$\mathbf{R} = \begin{bmatrix} \mathbf{S}_1 \mathbf{S}_1^T & \cdots & \mathbf{S}_1 \mathbf{S}_Q^T \\ \vdots & \ddots & \vdots \\ \mathbf{S}_Q \mathbf{S}_1^T & \cdots & \mathbf{S}_Q \mathbf{S}_Q^T \end{bmatrix} , \ \mathbf{P} = \begin{bmatrix} \mathbf{S}_1 \mathbf{S}_1^T & \cdots & \mathbf{O} \\ \vdots & \ddots & \vdots \\ \mathbf{O} & \cdots & \mathbf{S}_Q \mathbf{S}_Q^T \end{bmatrix} . \tag{5.8}$$

The eigenvectors corresponding to the eigenvalues sorted in a descending order and the largest eigenvalue are denoted by $\mathbf{w}_i^{(1)}, \mathbf{w}_i^{(2)}, \ldots, \mathbf{w}_i^{(\Sigma_i I_i)}$ and $\lambda^{(1)}$, respectively. The largest eigenvalue can be interpreted as the similarity of those multiple datasets and the eigenvectors corresponding to the large eigenvalues can be interpreted as the weight vectors that increase the correlation between these datasets.

In this chapter, MCCA is applied substituting the EEG signal of test data $\mathbf{X}$ for $\mathbf{S}_1$, the artificial reference signal $\mathbf{Y}$ for $\mathbf{S}_2$, and the training reference signal $\mathbf{Z}$ for $\mathbf{S}_3$. Applying MCCA to these datasets, spatial filters $\mathbf{w}_{\mathbf{X}}^{(l)}$ $(l = 1, \ldots, L \ll C + 2H + C)$ that are eigenvectors corresponding to top L eigenvalues and the largest eigenvalue $\lambda_k^{(1)}$ are obtained. With those spatial filters, the test data $\mathbf{X}$ and the training reference signal $\mathbf{Z}_k$ are projected. Subsequently, correlation coefficients between these two projections are obtained as:

$$\rho_k^{(l)} = \mathrm{Corr}\left(\left(\mathbf{w}_{\mathbf{X}}^{(l)}\right)^T \mathbf{X}, \left(\mathbf{w}_{\mathbf{X}}^{(l)}\right)^T \mathbf{Z}_k\right) . \tag{5.9}$$

The correlation coefficients are computed for each command k $(k = 1, \ldots, K)$ and each spatial filter $\mathbf{w}_{\mathbf{X}}^{(l)}$ $(l = 1, \ldots, L)$ following Equation (6.9). The largest eigenvalue and the correlation coefficients are combined and used for detecting a command as follows:

$$k^\star = \arg\max_k \ \lambda_k^{(1)} \sum_{l=1}^{L} \mathrm{sign}\left(\rho_k^{(l)}\right) \cdot \left(\rho_k^{(l)}\right)^2 . \tag{5.10}$$

It is worth noting that there is a difference between the proposed method and a method proposed by Zhang *et al.*for detecting the frequency of SSVEP using MCCA [96]. In Zhang *et al.*'s method, MCCA was employed to obtain an optimal reference signal instead of the artificial reference signal by extracting common features in the training set, while the proposed method

employs MCCA to extract common features between the test data, the artificial reference signal, and the training set.

5.2.6 Performance Criterion

To evaluate performance, accuracy of command detection was measured. In addition, simulated ITR [2] was also measured as follows:

$$\mathrm{ITR} = \frac{60}{U}\left[\log_2 K + P\log_2 P + (1-P)\log_2\frac{1-P}{K-1}\right], \tag{5.11}$$

where the unit of the ITR is bits/min, U [sec] is the mean time to input one command, K is the number of selectable commands (here, K = 16), and P is the accuracy. In this chapter, U was defined as $U = T_g + T_d + T_a$, where T_g [sec] was the time for the subjects to shift their gazes, T_d [sec] was the delay of SSVEP onset described before, and T_a [sec] was the length of the signal to analyze. Here these were set as $T_g = 0.5$, $T_d = 0.15$ and T_a was 0.1 to 2.0 sec with an interval of 0.1 sec, respectively. These values were set to simulate and evaluate an online performance.

5.3 Results and Discussions

5.3.1 Experimental results

Number of spatial filters

Figure 5.2 shows box plots of the distributions of the ITR analyzing 0.5 sec of the data length with respect to the number of spatial filters $L = 1, \ldots, 8$. The bands inside the boxes represent the median. Figure 5.2 illustrates that the ITR tended to increase with the number of spatial filters increased. The ITR

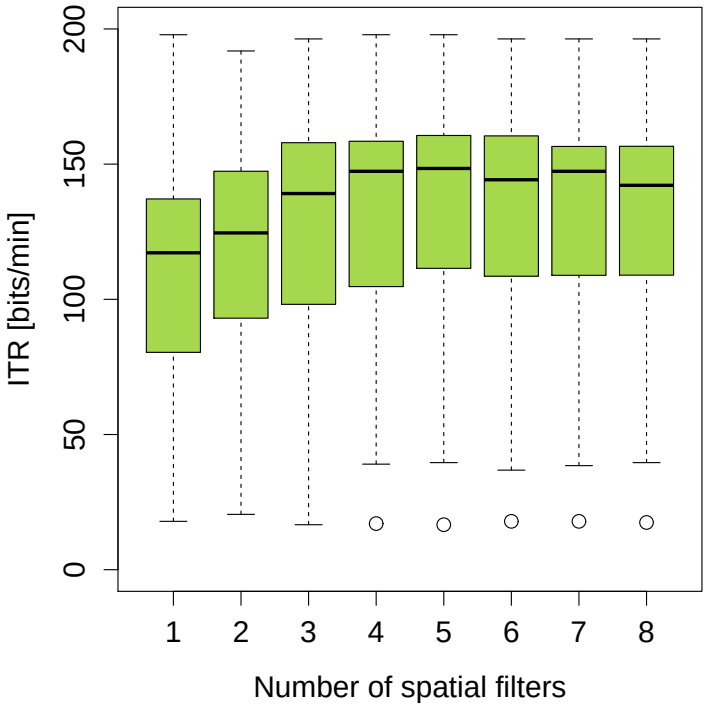


FIGURE 5.2: Box plot of the distribution of the ITR with respect to the number of spatial filters L.

remained steady in the case of using more than five spatial filters. These results suggest the efficacy of using more than one spatial filter, i.e., the weight vector corresponding to the largest eigenvalue.

Accuracy and ITR

Figures 5.3a and 5.3b show the averaged accuracy and ITR across the subjects with respect to the data length to be analyzed, respectively. The number of spatial filters in the proposed method was set to $L = 5$ according to the results previously described in Section 5.3.1.

As shown in Figure 5.3, the proposed method outperformed the contrast method in terms of both accuracy and ITR especially when the data length was shorter than approximately 1.0 sec. The proposed method achieved at the peak of averaged ITR with the data length of 0.5 sec while the contrast method achieved at the peak of averaged ITR with the data length of 0.7 sec.

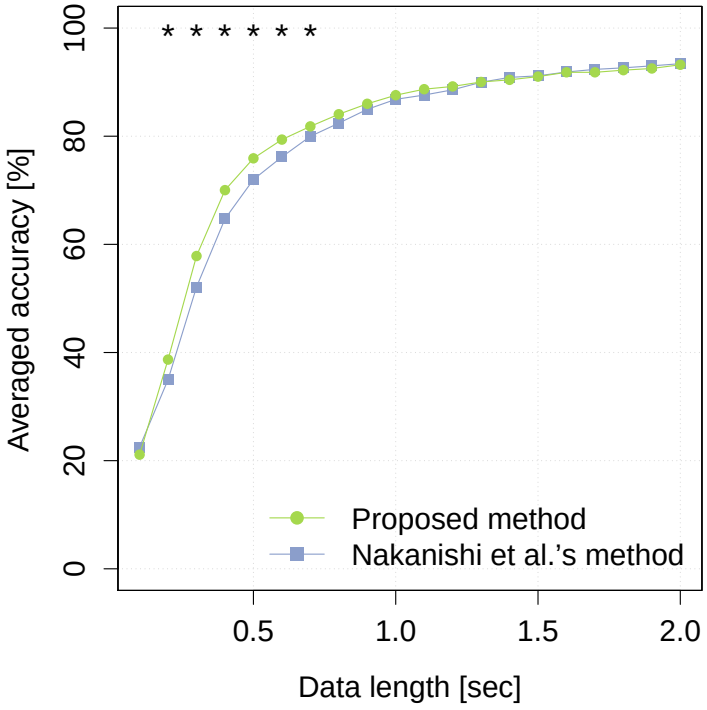
When the data length was longer than 1.0 sec, the proposed method and the contrast method showed similar performance.

5.3.2 Statistical verifications

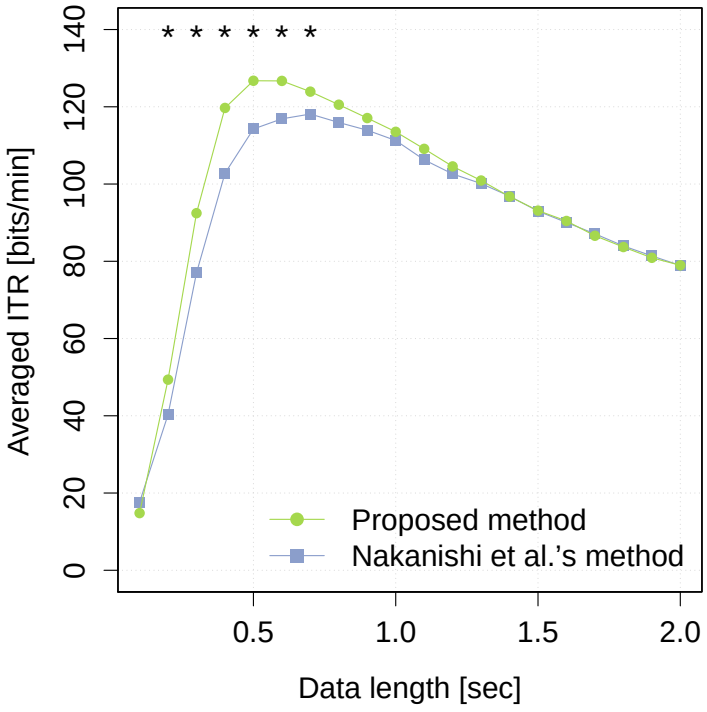
A paired-samples t-test was conducted to compare the accuracy and the ITR of the proposed method and the contrast method. The proposed method showed significantly higher accuracy and ITR than the contrast method in the case of using the data length of 0.2 to 0.7 sec ($p < 0.05$) as illustrated in Figure 5.3. For instance, in the case of using the data length of 0.5 sec, there was a significant difference in the accuracy of the proposed method ($75.91 \pm 22.03\%$) and the contrast method ($71.99 \pm 20.51\%$); $t(10) = 4.08,\ p < 0.01$, and in the ITR of the proposed method (126.74 ± 55.95 bits/min) and the contrast method (114.25 ± 51.94 bits/min); $t(10) = 4.10,\ p < 0.01$. Moreover, there was also a significant difference in the highest ITR of the proposed method (126.74 ± 55.95 bits/min; $T_a = 0.5$ sec) and the contrast method (118.12 ± 48.15 bits/min; $T_a = 0.7$ sec); $t(10) = 2.93,\ p < 0.05$. Taken together, these results suggest that the proposed method allowed us to detect commands faster and more accurate than the contrast method.

5.4 Conclusions

A novel method to recognize SSVEPs for mixed-coded SSVEP-based BMIs was proposed. The proposed method employed MCCA to obtain spatial filters which enhanced SSVEP components. The experimental results showed a significant improvement in command recognition accuracy and ITR using the proposed method. These results indicate the efficacy of employing MCCA to recognize the frequency and the phase of SSVEPs.



(A)



(B)

FIGURE 5.3: Averaged (a) accuracy and (b) ITR across the subjects with respect to the data length. The asterisks indicate significantly increased (a) accuracy and (b) ITR compared with Nakanishi *et al.*'s method.

Chapter 6

Asynchronous SSVEP-Based BMI

One of the challenges in the area of brain–computer interfacing (BMI) is to develop an *asynchronous* BMI or a *self-paced* BMI that detects whether a user intends to pass messages. This chapter proposes a novel asynchronous BMI that uses mixed frequency and phase-coded visual stimuli, which can provide high-speed and accurate command entries. The mixed-coded visual stimuli were presented as flickers with a following blank interval to synchronize the recorder of electroencephalogram (EEG) with the stimuli, which was aimed to detect the phase in an asynchronous situation. For decoding from the measured EEG, multiset canonical correlation analysis (MCCA) was efficiently exploited for recognizing the intentional state and the intending command. The proposed asynchronous BMI was tested on 11 healthy subjects. The proposed decoder was capable of discriminating between the IC/NC state and determining the command faster and more accurately than the contrast methods, achieving AUC of 0.9191 ± 0.1206 and command recognition accuracy of $91.08 \pm 13.97\%$ with data lengths of 3.0 seconds. The BMI based on mixed-coded visual stimuli was able to be implemented in an asynchronous manner, and the MCCA-based decoder outperformed the conventional ones in terms of discriminability of intentional states and command recognition accuracy. The present study showed that an asynchronous BMI can be implemented with mixed-coded visual stimuli for the first time, which enables a large increase in the number of choices/commands.

6.1 Introduction

A few works on decoders for asynchronous SSVEP-based BMIs have been reported in the last decade [43, 44, 41, 45, 46]. Recently developed methods for decoding the states and commands in the case of frequency-coded asynchronous BMIs include: *classification based on common spatial pattern (CCSP)* [44], *thresholding based on minimum energy combination (TMEC)* [41], *thresholding ratios of CCA coefficients (TRCC)* [46], and *clustering canonical correlations (CCC)* [100]. Parini *et al.*presented the CCSP method that employs common spatial pattern (CSP) to extract features and classify IC/NC states [44]. CSP can be used to find spatial filters that maximize the variance of one class and minimize the variance of the other class [101]. This method is intended to find a spatial filter that maximizes the variance of the IC class and minimizes that of the NC class. After the spatial filtering based on CSP, the method extracts features based on the periods, taking the advantage of the characteristics of SSVEPs, i.e., periodic signals. Lastly, a multi-class classifier is used to determine the IC/NC state and the command. Cecotti reported on the TMEC method [41] using a minimum energy combination (MEC) approach [102]. In the method, normalized frequency powers are computed for each stimulus frequency following an MEC approach, and the frequency is detected if the normalized frequency power is greater than a predefined threshold. The method also employs pseudo stimulus frequencies to improve its reliability. Xia *et al.*proposed the TRCC method [46] to discriminate between IC/NC states based on correlation coefficients computed by the standard canonical correlation analysis (CCA) [39]. With the standard CCA approach, correlation coefficients are computed for each stimulus frequency. The method classifies IC/NC states by a predefined threshold for a ratio of the second largest coefficient to the largest coefficient, which is based on an assumption that the ratio would be small in the IC state while it would be large in the NC state.

In regard to TMEC and TRCC, the threshold to determine the IC/NC state is determined empirically or adjusted manually, however, the optimization of the threshold is generally strenuous. Furthermore, in the CCC method presented by Poryzara and Materka [100], CCA is first applied to obtain correlation coefficients as features and feature vectors for each stimulus frequency are constructed. With the feature vector, centroids of the IC state and that of the NC state are obtained by k-means clustering analysis where the number of clusters is set to 2. Then, a new feature vector is classified based on the nearest neighbor method with those two centroids.

The main problems of the above-mentioned asynchronous BMIs are threefold: 1) to the knowledge of the authors, all the asynchronous BMIs based on SSVEP use *frequency*-coded visual stimuli, which have a limitation on the number of choices/commands being presented on a display; 2) it appears that the performance levels in detecting an intentional state and an intended command are not sufficiently high. Accordingly, to ensure the detection, the final decision of the states and the command is made if the detected commands are consecutive and coherent for the last few times; in other words, the detection procedure has to be repeatedly applied to determine the state and a command [44, 46]; 3) in design, the NC state can be detected by these methods. However, even though all the methods mentioned above developed BMIs that can detect the NC state, except for the TRCC, the experiment evaluated only the case where subjects were supposed to be in the IC state all the time. This means that there was no verification of the ability to discriminate between IC/NC states [44, 41, 100].

In this chapter, we propose to use not only frequency but also phase, known as mixed-coded visual stimuli [48, 37, 49] in an asynchronous BMI, yielding an increase of the number of choices/commands. So as to detect the phase of SSVEPs asynchronously, the visual stimuli were presented as flickers putting a non-flickering blank interval between each flickering interval

to synchronize the recorder of EEG with the stimuli. For discriminating between IC and NC states as well as detecting commands in the case of mixed-coded visual stimuli, the proposed system incorporates a novel method that exploits multiset canonical correlation analysis (MCCA) [47, 49], yielding spatial filters that enhance the SSVEP component. We hypothesized that the proposed implementation with the proposed MCCA decoder would increase the accuracy in command detection compared to the conventional decoders in the classical setting of frequency-coded stimuli.

The rest of the chapter is organized as follows: Section 6.2 describes the methods, including the conducted experiment and the proposed method to implement an asynchronous BMI based on mixed-coded visual stimuli; Section 6.3 summarizes the experimental results, comparing them to the conventional decoders; Section 6.4 describes the comprehensive evaluation of the performance of the proposed asynchronous BMI; Section 6.5 concludes the chapter and suggests future works.

6.2 Methods

6.2.1 Experimental Settings and Data Acquisition

Ten males and one female took part in our experiment as unpaid volunteers and were given informed consent forms. Their age ranged from 21 to 28 (mean age = 23.3 $\pm$ 2.0), and six of them had prior BMI experience. All subjects were healthy and had normal or corrected-to-normal vision. This study was approved by the research ethics committee of Tokyo University of Agriculture and Technology.

We used Ag/AgCl active electrodes, which are products of Guger Technologies (g.tec) named g.SCARABEO, g.SCARABEOgnd (for GND), and g.GAMMAearclip (for reference, earclip type) for recording EEG data. These were driven by a

g.GAMMAbox (g.tec) power supply unit. Fifteen electrodes were located at Pz, P1, P2, P3, P4, P5, P6, POz, PO3, PO4, PO7, PO8, Oz, O1, and O2 in accordance with the international 10–10 system [103]. The electrodes for GND and reference were AFz and A1, respectively. The signals were amplified by MEG-6116 (Nihon Kohden), which applied lowpass and highpass analog filters for each channel. The cutoff frequencies of the lowpass and the highpass filters were set to 100 Hz and 0.5 Hz, respectively. The EEG signal was sampled by an A/D converter (AIO-163202F-PE, Contec) with a sampling rate of $f_s = 1200$ Hz. The signals were recorded with the Data Acquisition Toolbox from the MATLAB (MathWorks). In an offline data analysis, all signals were first band-pass filtered between 5 and 50 Hz. Signals of 0.15 seconds just after visual stimuli started flickering were neglected and considered a latency of SSVEP [99].

A 24-inch LCD monitor (BenQ, XL2411T) with a resolution of 1,920 $\times$ 1,080 and a refresh rate of 120 Hz was used to display targets. The subjects were seated in a comfortable chair in front of a display 55 cm away so that they could look at the display straight ahead.

6.2.2 Mixed-coded Visual Stimuli

Figure 6.1 shows an image of displayed visual stimuli as targets. In the present study, a speller was implemented by presenting visual stimuli with letters. As illustrated in Figure 6.1, 28 visual stimuli with letters as targets were displayed, and a cross was also displayed at the center of the screen. The size of the visual stimuli was 4 cm on a side, and the interval between two adjacent targets was 2.5 cm. Visual stimuli were square and reversed to black and white according to the frequencies and phases of 13.0–16.5 Hz with an interval of 0.5 Hz and 0–270 degree with an interval of 90 degree. Each stimulus was presented with a specific combination of the frequency

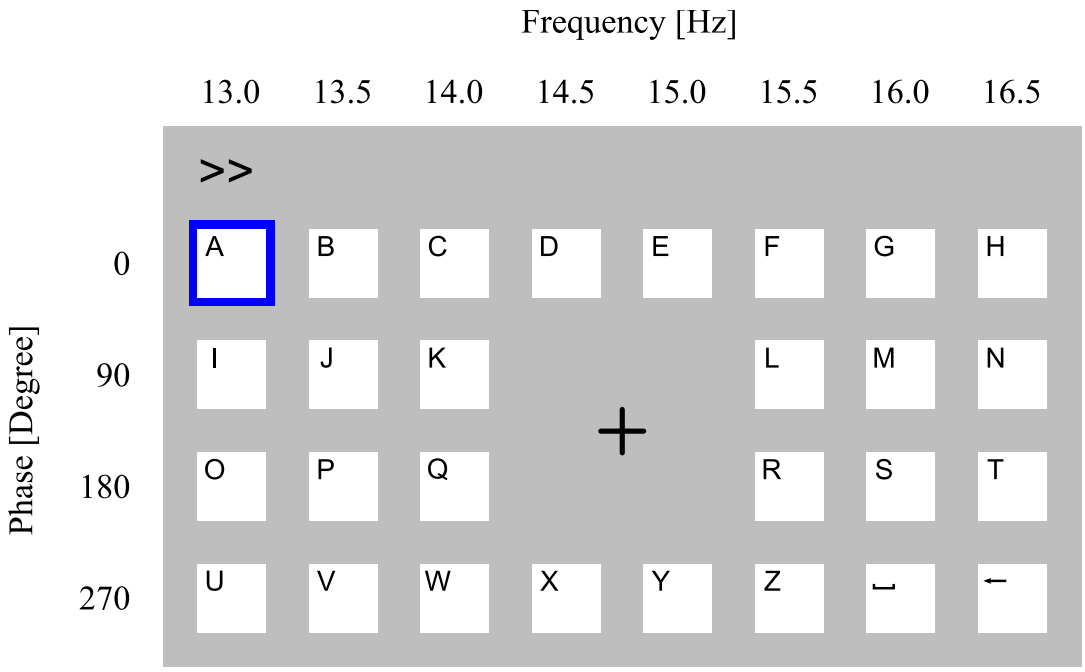


FIGURE 6.1: Image of displayed visual stimuli for the proposed asynchronous BMI. In the experiment, subjects were asked to gaze at a part framed by a blue rectangle. Each visual stimulus flickered according to the frequency and phase shown at the top and left of the image, respectively.

and phase, according to previous works on SSVEP-based BMIs using mixed frequency and phase coded visual stimuli [48, 37, 49], as shown in Figure 6.1. To produce those stimuli, an approximation method [92] was used. In this approach, a sequence of visual flickers $s(f, \phi, i)$ at the i-th frame with frequency f and phase ϕ was used as follows:

$$s(f, \phi, i) = \text{square}[2\pi f(i/\text{RefreshRate}) + \phi], \tag{6.1}$$

where square$[\cdot]$ denotes the operator that generates a 50% duty cycle square wave with levels 0 and 1 [92].

6.2.3 Design of an Asynchronous BMI

Figure 6.2 illustrates the time course of the data analysis in the proposed asynchronous BMI. The bottom of the figure shows an example of outputs.

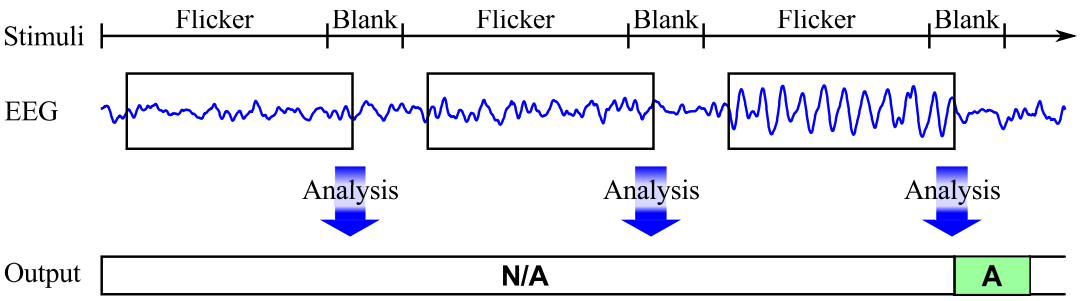


FIGURE 6.2: Time course of presenting visual stimuli and analyzing the EEG signal.

The proposed BMI repeatedly flickers the visual stimuli shown in Figure 6.1 for 3.0 seconds and stops flickering for 0.5 seconds as a blank interval. Every 3.0 seconds, a windowed EEG with a frame size of 0.5–3.0 seconds is analyzed and classified into either commands or NC by the approach described in Section 6.2.4. It should be emphasized that the underlying idea behind the blank interval of 0.5 seconds is to synchronize the recorder of EEG with the stimuli in order to decode the phase of SSVEPs.

6.2.4 State and Command Detection

A key to implementing a mixed-coded BMI is a method for detecting commands. A recently developed decoder by Nakanishi *et al.*for a synchronous BMI based on mixed-coded visual stimuli [37] uses the standard CCA [39] together with a reference signal composed of a training data set. This synchronous BMI showed the highest ITR ever reported at that time. Motivated by the method above, in this chapter, we employed a simpler command detector based on MCCA [49] and a support vector machine (SVM) [104]. The proposed MCCA-based decoder would be more efficient than Nakanishi *et*

al.'s approach, since the proposed decoder only needs to solve a single eigenvalue problem, though the method [37] should solve three different eigenvalue problems. Figure 6.3 illustrates the architecture of the proposed decoder. As shown in Figure 6.3, MCCA is first applied to three datasets following test data $\mathbf{X}$, an artificial reference signal $\mathbf{Y}_k$, and a training reference signal $\mathbf{Z}_k$ which is composed by utilizing a training set. Thereby, correlation coefficients are computed to obtain features. Those features are classified as IC$_k$ $(k = 1, \ldots, K)$ state or NC state using multi-class SVM, where k denotes an index of a command assigned to the kth stimulus. Details for each part are described below.

Feature Extraction Using Multiset CCA

MCCA is a statistical method to evaluate a similarity between multiple datasets [47, 95]. In the proposed decoder, MCCA was first applied to three datasets: $\mathbf{X}$, $\mathbf{Y}_k$, and $\mathbf{Z}_k$. Using spatial filters with respect to the test data obtained by MCCA, correlation coefficients between linear combinations of the test data and that of the training reference signal were computed and used as features.

MCCA is a generalization of CCA to multiple datasets [47, 95]. It finds spatial filters that maximize a correlation between weighted linear combinations of each dataset, called canonical variates. Let $\mathbf{S}_i \in \mathbb{R}^{I_i \times N}$ $(i = 1, \ldots, Q)$ be Q multi-channel datasets that were normalized to have zero mean and unit variance. Their linear combinations, canonical variates, are denoted by $\mathbf{w}_i^T \mathbf{S}_i$, where $\mathbf{w}_i \in \mathbb{R}^{I_i \times 1}$ is a spatial filter. With the MAXVAR criterion [47, 94, 97], MCCA solves the following maximization problem to find spatial filters

$\mathbf{w} = [\mathbf{w}_1^T, \mathbf{w}_2^T, \ldots, \mathbf{w}_Q^T]^T$:

$$\rho = \max_{\mathbf{w}_1,\ldots\mathbf{w}_Q} \sum_{i \neq j}^{Q} \mathbf{w}_i^T \mathbf{S}_i \mathbf{S}_j^T \mathbf{w}_j \tag{6.2}$$

$$\text{s.t.} \quad \frac{1}{Q} \sum_{i=1}^{Q} \mathbf{w}_i^T \mathbf{S}_i \mathbf{S}_i^T \mathbf{w}_i = 1. \tag{6.3}$$

Using a Lagrange multiplier technique, this objective function can be transformed into a generalized eigenvalue problem:

$$(\mathbf{R} - \mathbf{P})\mathbf{w} = \rho \mathbf{P}\mathbf{w}, \tag{6.4}$$

where

$$\mathbf{R} = \begin{bmatrix} \mathbf{S}_1\mathbf{S}_1^T & \cdots & \mathbf{S}_1\mathbf{S}_Q^T \\ \vdots & \ddots & \vdots \\ \mathbf{S}_Q\mathbf{S}_1^T & \cdots & \mathbf{S}_Q\mathbf{S}_Q^T \end{bmatrix}, \ \mathbf{P} = \begin{bmatrix} \mathbf{S}_1\mathbf{S}_1^T & \cdots & \mathbf{O} \\ \vdots & \ddots & \vdots \\ \mathbf{O} & \cdots & \mathbf{S}_Q\mathbf{S}_Q^T \end{bmatrix}. \tag{6.5}$$

The eigenvectors corresponding to the eigenvalues are sorted in a descending order, and the largest eigenvalue are denoted by $\mathbf{w}_i^{(1)}, \mathbf{w}_i^{(2)}, \ldots, \mathbf{w}_i^{(\Sigma_i I_i)}$ and $\lambda^{(1)}$, respectively. The largest eigenvalue can be interpreted as representing the similarity of those multiple datasets, and the eigenvectors corresponding to the large eigenvalues can be interpreted as representing the spatial filters that increase the correlation between these datasets.

In the proposed architecture, $\mathbf{S}_1$, $\mathbf{S}_2$, and $\mathbf{S}_3$ correspond to the EEG signal of test data $\mathbf{X}$, the artificial reference signal $\mathbf{Y}_k$, and the training reference signal $\mathbf{Z}_k$. The artificial reference signal $\mathbf{Y}_k$ with respect to a stimulus frequency f_k consisted of a Fourier series of simulated stimulus signals given as

$$\mathbf{Y}_k = \begin{bmatrix} \mathbf{Y}_k(0) & \mathbf{Y}_k(1/f_s) & \cdots & \mathbf{Y}_k((N-1)/f_s) \end{bmatrix}, \tag{6.6}$$

where

$$\mathbf{Y}_k(t) = \Big[\sin(2\pi f_k t) \quad \cos(2\pi f_k t) \quad \cdots \quad \sin(2H\pi f_k t) \quad \cos(2H\pi f_k t)\Big]^T, \tag{6.7}$$

$$\mathbf{Y}_k = [\sin(2\pi f_k t), \cos(2\pi f_k t), \ldots, \sin(2H\pi f_k t), \cos(2H\pi f_k t)]^T, \tag{6.8}$$

where H is the number of harmonics in the Fourier series, the first two components are the sinusoids of the fundamental frequency f_k, and the others are harmonics. Additionally, the training reference signal $\mathbf{Z}_k$ corresponding to the kth stimulus was defined as the averaged EEG signals of a training set across trials. Applying MCCA to these datasets, the spatial filters $\mathbf{w}_{\mathbf{X}}^{(l)}$ $(l = 1, \ldots, L \ll C + 2H + C)$ that were eigenvectors corresponding to the top L eigenvalues and the largest eigenvalue $\lambda_k^{(1)}$ were obtained. With those spatial filters, the test data $\mathbf{X}$ and the training reference signal $\mathbf{Z}_k$ were projected. Subsequently, correlation coefficients between these two projections were obtained as follows:

$$\rho_k^{(l)} = \mathrm{Corr}\left(\left(\mathbf{w}_{\mathbf{X}}^{(l)}\right)^T \mathbf{X}, \left(\mathbf{w}_{\mathbf{X}}^{(l)}\right)^T \mathbf{Z}_k\right). \tag{6.9}$$

The correlation coefficients were computed for each command k $(k = 1, \ldots, K)$ and each spatial filter $\mathbf{w}_{\mathbf{X}}^{(l)}$ $(l = 1, \ldots, L)$ following Equation (6.9). The largest eigenvalue and the correlation coefficients were combined as

$$r_k^{(l)} = \lambda_k^{(1)} \mathrm{sign}\left(\rho_k^{(l)}\right) \cdot \left(\rho_k^{(l)}\right)^2 \tag{6.10}$$

and used as features. The following feature vector

$$\mathbf{u} = \begin{bmatrix} r_1^{(1)} & \cdots & r_k^{(l)} & \cdots & r_K^{(L)} \end{bmatrix} \in \mathbb{R}^{1 \times K \cdot L} \tag{6.11}$$

was used to estimate the intentional state and the intended command.

Classification Using Multi-class SVM

To estimate the intentional state and the command simultaneously, multi-class classification was employed. Specifically, M-class SVM with one-versus-rest approach [104] was used, where $M = K + 1$ indicates the total number of classes to classify, that is, the K classes of the IC state and the NC class. In the one-versus-rest approach, M independent SVMs were utilized. For the SVMs, a radial basis function (RBF) kernel was employed, and the parameters were set to $c = 1$, $\gamma = \frac{1}{K \cdot L}$, where c is a cost parameter, γ is a parameter of the RBF kernel, and $K \cdot L$ indicates the dimension of the feature vector defined in Equation (6.11).

6.2.5 Experiment

The subjects performed the following tasks. In a trial, the subjects gazed at one of the visual stimuli or the center cross that was instructed by a blue rectangle in a random order. Trials corresponding to each stimulus were referred to as IC$_k$ $(k = 1, \ldots, 28)$ trials, and trials corresponding to the center cross were referred to as NC trials. One trial consisted of 3.0 seconds of flickering time and 0.5 seconds of blank interval. It corresponds to a set of the flickering interval and the blank interval in Figure 6.2. First, non-flickering targets and a blue rectangle were displayed. In flickering time, all stimuli started flickering and the subjects gazed at the instructed target, that is, one of the visual stimuli or the center cross. In the blank interval, all stimuli stopped flickering and the blue rectangle instructing the next target was displayed.

IC_k $(k = 1, \ldots, 28)$ trials were performed 10 times each, and NC trials were performed 20 times; thus, 280 IC trials and 20 NC trials equal 300 trials in total that were performed by one subject. After every 10 trials, the task was stopped in order to reduce visual fatigue, and it was restarted when the subjects pressed the space bar.

6.2.6 Contrast Methods

The conventional decoders for asynchronous SSVEP-BMI described in Section 6.1 were employed as contrast decoders to evaluate the validity of the proposed decoder. The contrast methods included CCSP [44], TMEC [41], TRCC [46], and CCC [100]. CCSP and TRCC originally identify the command by using consecutive frames, and only when some consecutive and coherent commands were detected the final output of command was determined. To fairly and simply compare the performances of the proposed decoder and the contrast ones, these methods were implemented to determine the command/state based on a single detection.

The parameters of these methods were set as follows:

CCSP Multi-class regularized discriminant analysis (RDA) [105] with a boosting technique [106] was employed as a classifier. Two parameters for RDA were selected using a grid search procedure based on 4-fold cross-validation. Moreover, for a boosting technique, five weak classifiers were trained to create a strong classifier.

TMEC The parameter of the softmax function α was set to 0.25 and the threshold ζ, which discriminated between IC/NC states, was set to 0.5. These parameters were set according to [41]. In addition, the pseudo stimulus frequencies were chosen as 13.25–16.25 Hz with an interval of 0.5 Hz, which were averages of the two adjacent frequencies used to present visual stimuli.

TRCC The threshold θ, which discriminated between IC/NC states, was chosen according to the Bayesian classification rule using a training set. It set the threshold θ where the posterior probability of the IC state was equal to that of the NC state.

CCC The threshold δ of Euclidean distance between two centroids of IC/NC states was set to 0.25, according to [100]. If the distance between those two centroids did not achieve the threshold, two centroids that had the maximum distance were chosen.

6.2.7 Performance Evaluation

The performance of these methods were evaluated based on 5-fold cross-validation. As a criterion of the performance regarding with discriminability of the IC/NC states, area under the curve (AUC) was measured with respect to receiver operating characteristic (ROC) curve plotted sensitivity against $1-$ specificity varying a threshold. The sensitivity (SEN) and specificity (SPC) were defined as follows:

$$\mathrm{SEN} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}, \tag{6.12}$$

$$\mathrm{SPC} = \frac{\mathrm{TN}}{\mathrm{FP}+\mathrm{TN}}, \tag{6.13}$$

where TP (true positive) is the number of correctly detected IC trials, FN (false negative) is the number of falsely detected IC trials, FP (false positive) is the number of falsely detected NC trials, and TN (true negative) is the number of correctly detected NC trials. For the proposed method and the contrast methods, the parameter ϵ in the following equations was set at various values, respectively:

CCSP $\Delta_{\mathrm{NC}} + \epsilon$, where Δ_{NC} was a confidence score corresponding to NC state;

TMEC $\zeta + \epsilon$, where ζ was the predefined threshold;

TRCC $\theta + \epsilon$, where θ was the fixed threshold which was determined using a training set;

CCC $d(\mathbf{u}_k^{\mathrm{new}}, \mathbf{C}_k^{\mathrm{NC}}) + \epsilon$, where $d(\mathbf{u}_k^{\mathrm{new}}, \mathbf{C}_k^{\mathrm{NC}})$ was a distance between a feature vector of the test data and the centroid of the NC state;

Proposed method $\Pi_{\mathrm{NC}} + \epsilon$, where Π_{NC} was a confidence score corresponding to the NC state.

Moreover, the following three types of command recognition accuracy were also measured: frequency (8 commands), phase (4 commands), and mixed frequency and phase (28 commands), denoted by FACC for frequency, PACC for phase, and MACC for mixed recognition accuracy, respectively. The command recognition accuracy was measured under the condition of TP, that is, when the IC trial was correctly detected as IC state.

After measuring their performances, statistical tests including Friedman's one-way analysis of variance (ANOVA) and *post hoc* tests were performed for AUC and FACC, and for each data length of EEG signal. These statistical tests were not performed on PACC and MACC, since the contrast decoders have not been developed for phase recognition. Friedman's ANOVA, which is a non-parametric test based on ranked data [107], was performed for the performance to detect differences in the methods, including the proposed decoder and the contrast ones. If the effect of the methods from Friedman's ANOVA was significant with a significance level of 0.05, *post hoc* tests were performed between the proposed decoder and each contrast decoder, according to the rank-based approach described by Siegel and Castellan [108]. A

TABLE 6.1: AUC using the data length of 3.0 seconds. The highest value for each subject and the average across the subjects are in bold.

Subjects	CCSP [44]	TMEC [41]	TRCC [46]	CCC [100]	Proposed method $L=1$	Proposed method $L=2$	Proposed method $L=3$
sub1	0.942	0.824	0.982	**0.994**	0.975	0.983	0.979
sub2	0.997	0.949	0.996	**0.999**	0.989	0.980	0.981
sub3	0.929	0.781	0.742	**1.000**	0.997	0.998	0.998
sub4	0.940	0.862	0.991	0.997	0.997	**1.000**	**1.000**
sub5	0.764	0.666	0.780	**0.937**	0.899	0.876	0.834
sub6	0.925	0.704	0.896	0.949	**0.950**	0.904	0.893
sub7	0.718	0.522	0.678	**0.855**	0.619	0.644	0.621
sub8	0.861	0.654	0.984	**0.998**	0.996	0.996	0.996
sub9	0.992	0.904	0.959	**0.997**	0.982	0.975	0.978
sub10	0.842	0.623	0.842	0.845	**0.940**	0.914	0.873
sub11	0.678	0.470	0.656	**0.780**	0.766	0.704	0.612
Ave.	0.8716	0.7235	0.8642	**0.9409**	0.9191	0.9068	0.8878
S.D.	0.1095	0.1549	0.1312	0.0785	0.1206	0.1233	0.1455

significance level for the *post hoc* tests was also set to 0.05. The details of these analyses can be found in [107, 108, 109].

6.3 Results

6.3.1 Area Under the Curve (AUC)

Figure 6.4 shows the result of AUC which represents the classification accuracy of the IC/NC states using the proposed MCCA-based decoder and the contrast decoders. Table 6.1 also presents the result of AUC using the data length of 3.0 seconds. The number of spatial filters used in the proposed decoder was set to $L = 1, 2$, and 3.

As shown in Figure 6.4a, all the methods showed higher AUC with a longer data length. The averaged AUC of the proposed decoder achieved 0.9 with data lengths of 2.6 seconds ($L = 1$), 2.8 seconds ($L = 2$), respectively. Meanwhile, the averaged AUC of CCC [100] achieved 0.9 with a data

length of 2.2 seconds. On the contrary, TRCC [46], TMEC [41] and CCSP [44] did not achieve 0.9 of AUC. The proposed decoder achieved higher AUC with the smaller number of spatial filters, as shown in Table 6.1. Hence, in the following statistical tests, the results of the proposed decoder when the number of filters was set to $L = 1$ was employed as a representative of the proposed decoder. Reporting on the results of Friedman's ANOVA, there was a significant main effect of the methods on AUC for all data lengths of the EEG signal ($p < 0.05$). Thus, the *post hoc* tests were performed for each and all data lengths. The results of the *post hoc* tests are shown in Figure 6.4a by filled/empty markers for each contrast decoder. If the proposed decoder had significantly higher AUC ($p < 0.05$) compared to a contrast decoder, the marker on a line of the contrast decoder is filled, thus, the empty markers show no significant difference. For instance, from Figure 6.4a, we can see that the proposed decoder had significantly higher AUC than CCC when the data length was 0.6 seconds. It is worth noting that none of the contrast decoders showed significantly higher AUC than the proposed one. The proposed decoder based on MCCA showed significantly higher AUC compared to CCSP, TMEC, and TRCC especially when the data length was short, for example, when the data length was 1.0 second. There was no significant difference between the proposed decoder and CCC except for the case where the data length was 0.6 seconds.

6.3.2 Frequency Recognition Accuracy

Figure 6.5 shows the results of the frequency recognition accuracy (FACC) and Table 6.2 also shows it when the data length was 3.0 seconds. FACC serves as a criterion of how accurate the frequency of the detected command agreed with that of the correct command.

TABLE 6.2: Frequency recognition accuracy [%] using the data length of 3.0 seconds. The highest value for each subject and the average across the subjects are in bold.

Subjects	CCSP [44]	TMEC [41]	TRCC [46]	CCC [100]	Proposed method $L=1$	Proposed method $L=2$	Proposed method $L=3$
sub1	92.9	94.2	**100.0**	**100.0**	99.6	**100.0**	**100.0**
sub2	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
sub3	94.9	94.2	**100.0**	95.3	**100.0**	**100.0**	**100.0**
sub4	99.1	98.8	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
sub5	60.9	63.4	97.9	97.6	99.6	**100.0**	**100.0**
sub6	90.5	70.6	94.0	93.3	99.4	99.3	**100.0**
sub7	40.2	51.5	**88.6**	87.1	72.3	80.9	72.6
sub8	80.6	91.1	**100.0**	99.6	**100.0**	**100.0**	**100.0**
sub9	**100.0**	99.1	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
sub10	73.1	79.3	98.5	95.9	98.3	**99.5**	98.5
sub11	53.5	36.3	**90.1**	78.5	86.5	78.6	75.9
Ave.	80.52	79.87	**97.19**	95.20	95.97	96.20	95.19
S.D.	20.87	21.76	4.29	6.84	8.81	8.17	10.38

As shown in Figure 6.5a, all the methods showed higher frequency recognition accuracy with a longer data length in the same way as the AUC. FACC of the proposed decoder was slightly higher with a larger number of spatial filters, especially when the data length was as short as around 0.5–1.0 seconds. The averaged FACC of the proposed decoder achieved 90% with the data length of 1.8 seconds ($L = 1, 2$, and 3). Meanwhile, the averaged FACC of TRCC [46] and CCC [100] achieved 90% with the data length of 1.6 and 2.2 seconds, respectively. Conversely, TMEC [41] and CCSP [44] did not achieve 90% of accuracy. Although the proposed decoder showed slightly higher FACC with a larger number of spatial filters, in order to put importance on the discriminability of the IC/NC state, the results of the proposed decoder when the number of filters was set to $L = 1$, which achieved the highest AUC, was again employed as a representative of the proposed decoder in the statistical tests. Reporting on the results of Friedman's ANOVA, there was a significant main effect of the methods on FACC for all data lengths of the EEG signal ($p < 0.05$) as in the result of AUC. Thus, the *post hoc* tests

TABLE 6.3: Phase recognition accuracy [%] using the data length of 3.0 seconds. The highest value for each subject and the average across the subjects are in bold.

Subjects	Proposed method		
	$L = 1$	$L = 2$	$L = 3$
sub1	98.7	98.9	**100.0**
sub2	98.9	**99.2**	**99.2**
sub3	99.2	**99.6**	99.5
sub4	96.1	96.1	**96.8**
sub5	91.9	**95.7**	95.1
sub6	95.6	**98.2**	97.5
sub7	74.6	**76.6**	70.6
sub8	98.9	98.9	**99.2**
sub9	**99.2**	**99.2**	99.1
sub10	**98.7**	97.2	94.8
sub11	**66.6**	63.0	55.8
Ave.	92.57	**92.95**	91.61
S.D.	11.24	11.92	14.52

were performed for each and all data lengths. The results of the *post hoc* tests are shown in Figure 6.5a in the same way as Figure 6.4a. In particular, when the data length was as short as around 0.5–1.0 seconds, the proposed decoder showed significantly higher FACC than CCSP, TMEC, and CCC, whereas there was no significant difference between the proposed decoder and TRCC. It is worth noting that none of the contrast decoders showed significantly higher FACC than the proposed decoder.

6.3.3 Phase Recognition Accuracy

Figure 6.6 shows the results of the phase recognition accuracy (PACC) and Table 6.3 also shows it when the data length was 3.0 seconds. The proposed decoder showed slightly higher PACC with a longer data length and the larger number of spatial filters in the same way as FACC. The averaged PACC of the proposed decoder achieved 90% with the data length of 2.4 seconds ($L = 1$), 2.2 seconds ($L = 2$), and 2.1 seconds ($L = 3$), respectively.

TABLE 6.4: Mixed frequency and phase recognition accuracy [%] using the data length of 3.0 seconds. The highest value for each subject and the average across the subjects are in bold.

Subjects	Proposed method		
	$L = 1$	$L = 2$	$L = 3$
sub1	98.7	98.9	**100.0**
sub2	98.9	**99.2**	**99.2**
sub3	99.2	**99.6**	99.5
sub4	96.1	96.1	**96.8**
sub5	91.9	**95.7**	95.1
sub6	95.6	**97.5**	**97.5**
sub7	64.1	**69.5**	64.3
sub8	98.9	98.9	**99.2**
sub9	**99.2**	**99.2**	99.1
sub10	**97.1**	96.6	94.8
sub11	**62.3**	56.5	49.6
Ave.	91.08	**91.61**	90.46
S.D.	13.97	14.50	16.99

6.3.4 Mixed Frequency and Phase Recognition Accuracy

Figure 6.7 shows the results of the mixed frequency and phase recognition accuracy (MACC) and Table 6.4 also shows it when the data length was 3.0 seconds. The proposed decoder showed higher MACC with a longer data length and the larger number of spatial filters in the same way as FACC and PACC. The averaged mixed frequency and phase recognition accuracy of the proposed decoder achieved 90% with the data length of 2.8 seconds ($L = 1$) and 2.5 seconds ($L = 2, 3$), respectively.

6.4 Discussions

The proposed decoder achieved significantly higher AUC than TRCC [46], CCSP [44], and TMEC [41], as shown in Figure 6.4a. Correspondingly, it showed significantly higher FACC than CCC [100], CCSP, and TMEC, as seen in Figure 6.5a. Concerning the speed of the system, the proposed decoder

achieved higher AUC and FACC than the contrast ones, especially when the data length was as short as around 0.5–1.5 seconds, as can be seen in Figures 6.4a and 6.5a. Comprehensively taking account into the results of AUC and FACC, the experimental results indicate that the proposed decoder is capable of discriminating between the IC/NC state and detecting the frequency of SSVEPs faster and more accurately than the contrast decoders. Furthermore, it was shown that the proposed approach was able to detect the phase as well as the frequency of SSVEPs, which leads to the fact that the present study proved that the proposed decoder can possibly be applied to a mixed-coding based asynchronous BMI.

As can be seen in Figure 6.4b and Table 6.1, one of the subjects, sub7, had much lower AUC compared to the other subjects with the proposed decoder. Besides, sub7 showed low AUC almost regardless of the data length with the proposed decoder and TMEC, while he/she achieved higher AUC with longer data length when using CCSP, TRCC, and CCC. Such a phenomenon, which is referred to as "BMI-inefficiency," can be occasionally caused depending on several different factors, such as individual characteristics of a user, characteristics of classification algorithms of a BMI [110]. In this case, the poor performance may have been due to the data length of EEG signal used in the analysis. The proposed decoder and TMCC could have shown high AUC if the data length was long enough. These results provide an insight that a proper data length for analysis would be characteristic for each user.

There are several EEG features which have been utilized in BMIs. Besides SSVEP, ERP and ERD have also been widely used in asynchronous BMIs [111, 112, 113]. These BMIs have their own advantages and disadvantages [114] in common with synchronous and asynchronous BMIs. Generally speaking, BMIs based on ERD can be implemented without any external stimuli and are eye-gaze independent, yet they demand more user training compared to

the other BMIs. BMIs based on ERP, specifically ones based on P300 typically have a high accuracy and the number of choices/commands is not limited, yet tend to take a longer time as the number of choices/commands increases, where P300 is one of the ERP components and is elicited by infrequent (oddball) stimuli [115]. As described in Section 6.1, BMIs based on SSVEP likely to be faster and more accurate in terms of command recognition compared to the other BMIs. Although, the flickering visual stimuli can cause visual fatigue and a risk of triggering seizures. Furthermore, some asynchronous BMIs have been implemented as hybrid BMIs by efficiently combining two out of SSVEP, ERP, and ERD [116, 117, 118]. For instance, Panicker *et al.* introduced a hybrid asynchronous BMI utilizing SSVEP for discriminating the IC/NC states and P300 for command entries, and reported the averaged AUC of 0.859 [116]. In comparison with the other asynchronous BMIs [111, 112, 113, 115, 116, 117, 118], one of the biggest advantages of the proposed asynchronous BMI is a fast and accurate command recognition as in the synchronous BMIs based on SSVEP [26]. It achieved the averaged recognition accuracy of 91.08% with 28 commands using the data length of 3.0 seconds when the number of spatial filters was set to $L = 1$. Since there has not yet been any established or commonly-used criterion to evaluate the capability of discriminating the IC/NC states for asynchronous BMIs, it is difficult to directly compare the asynchronous performance. For all that, the averaged AUC of 0.919 of the proposed asynchronous BMI would be a considerable development in this research field. Furthermore, in several asynchronous BMIs [41, 46, 116], parameters, such as a threshold to discriminate the IC/NC states are required to be determined empirically and/or adjusted manually. Some of these approaches do not need user's training. On the other hand, the proposed decoder utilized the training data. As a result, the experimental results showed that the proposed decoder achieved sufficiently high performance without any additional effort of parameter tuning.

6.5 Conclusion

The present study proposed a novel asynchronous SSVEP-BMI based on mixed frequency and phase-coded visual stimuli. The proposed system presented the mixed-coded visual stimuli as flickers with a following blank interval to synchronize the recorder of EEG with the stimuli, in order to decode the phase of SSVEPs asynchronously. For feature extraction to recognize intentional (IC) or non-control (NC) states, as well as the intended command, multiset canonical correlation analysis (MCCA) based approach was exploited. Experimental results showed that the proposed decoder had better performance than the conventional ones in terms of discriminability of IC/NC states and the command recognition accuracy. Moreover, mixed frequency and phase-coding was first applied to an asynchronous SSVEP-BMI in the study, which will be a significant contribution to the progress of the asynchronous SSVEP-BMIs since it leads to a large increase of the number of choices/commands.

The most important limitation lies in the fact that the MCCA-based decoder requires the training data. In order to reduce user's training, it would be fruitful to validate the efficacy of adaptive learning to obtain an adequate training reference signal. Besides, there have been some attempts to utilize stimulus frequencies in a high-frequency range for BMIs [119] to reduce visual fatigue and avoid a risk to provoke epileptic seizures. Thus, it is worthwhile to develop the asynchronous BMIs to use the high stimulus frequencies. Finally, the scope of the present study was limited in terms of the evaluation in the offline analysis. Therefore, further work could assess the usability and performance of the proposed asynchronous BMI in an online experiment with a feedback.

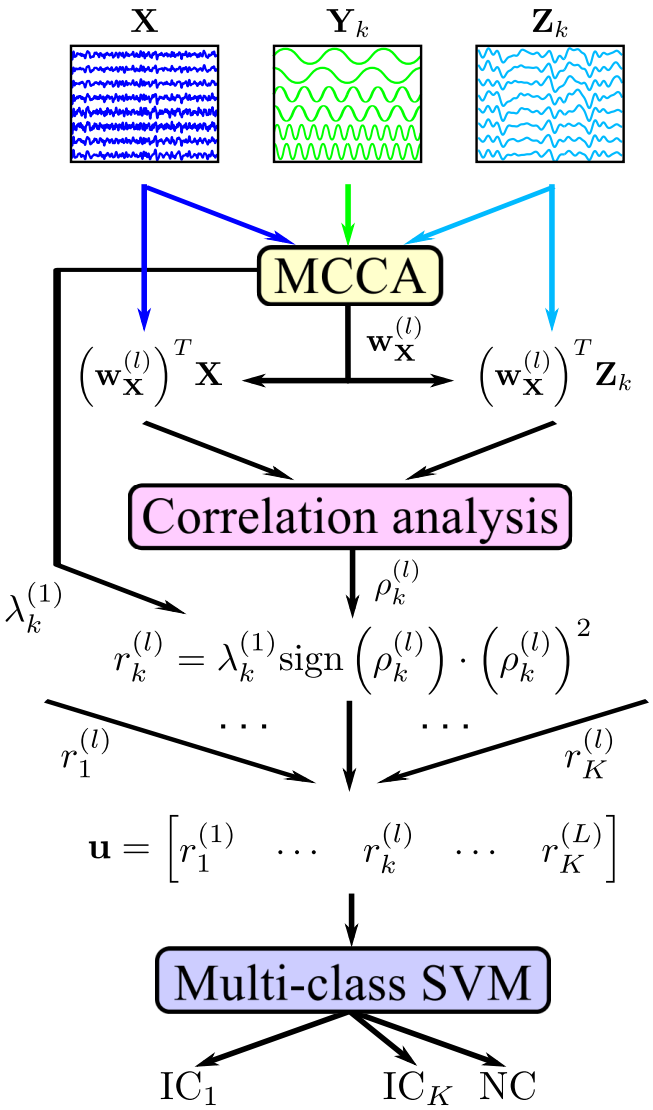


FIGURE 6.3: Architecture of the proposed decoder. First, correlation coefficients are obtained as features using MCCA. Next, those features are classified as IC_k or NC state using multi-class SVM, where $k = 1, \ldots, K$ indicates the index of commands.

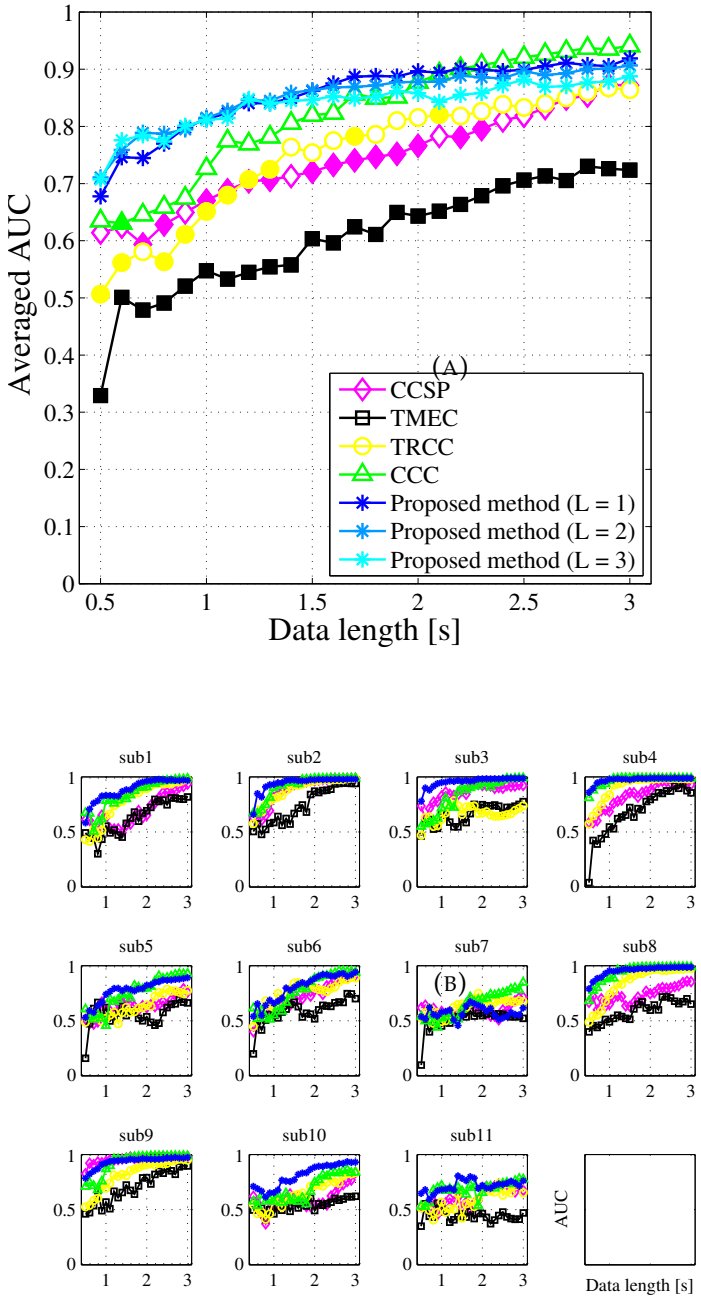


FIGURE 6.4: (a) Averaged AUC across the subjects and (b) that of each individual subject with respect to the data length using the proposed decoder and the contrast ones, respectively. The legend of (b) is consistent with the one in (a). In (a), filled markers represent significant differences comparing the results of the proposed decoder and each contrast decoder. In (b), the result of the proposed decoder when L was set to 1 is only shown as the representative of the proposed decoder for the sake of visibility.

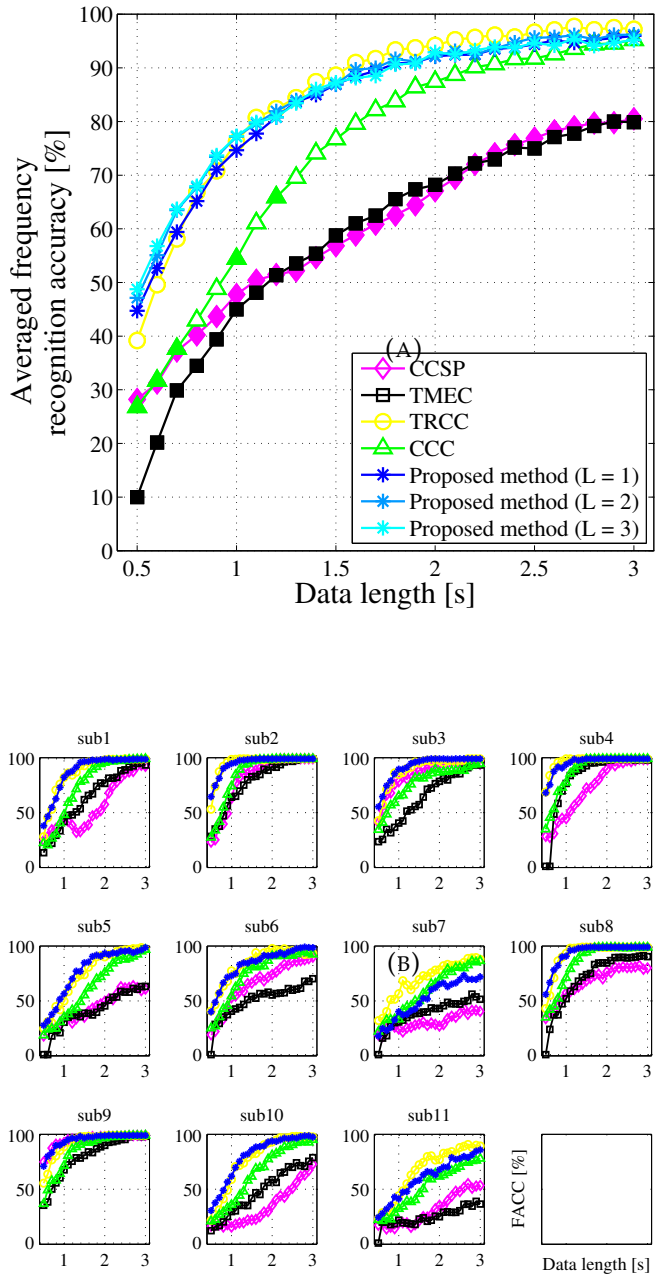


FIGURE 6.5: (a) Averaged frequency recognition accuracy across the subjects and (b) that of each individual subject with respect to the data length using the proposed decoder and the contrast ones, respectively. The legend of (b) is consistent with the one in (a). In (a), filled markers represent significant differences comparing the results of the proposed decoder and each contrast decoder. In (b), the result of the proposed decoder when L was set to 1 is only shown as the representative of the proposed decoder for the sake of visibility.

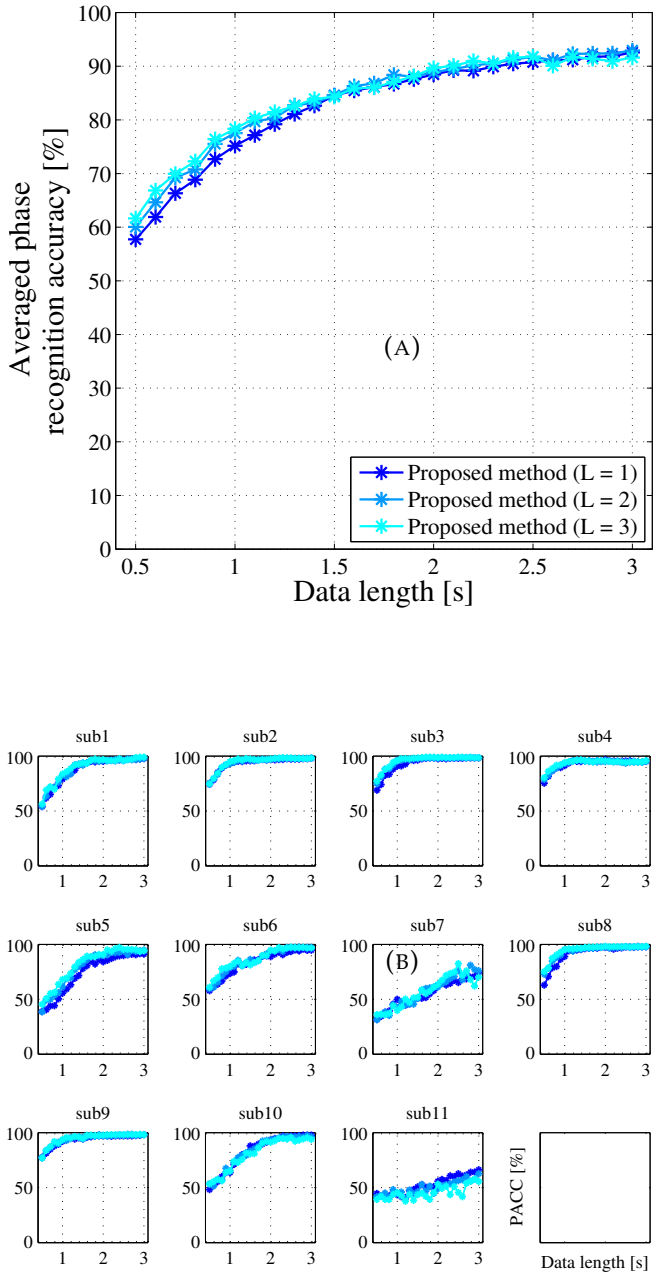


FIGURE 6.6: (a) Averaged phase recognition accuracy across the subjects and (b) that of each individual subject with respect to the data length using the proposed decoder, respectively. The legend of (b) is consistent with the one in (a).

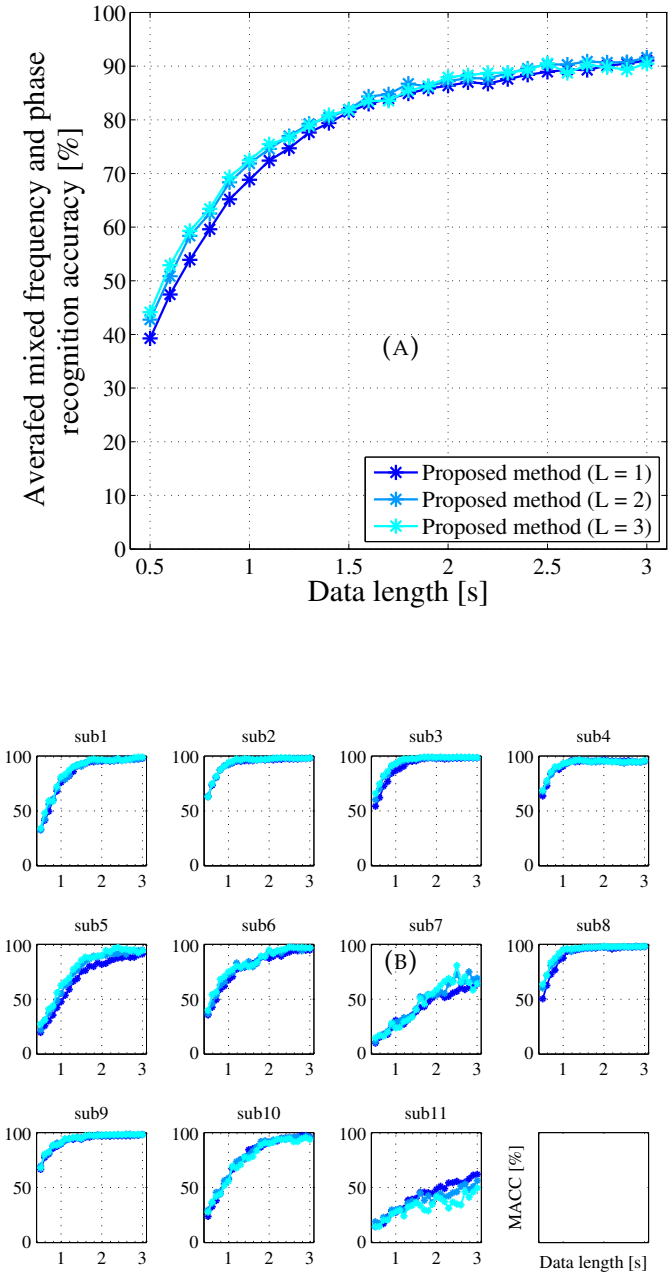


FIGURE 6.7: (a) Averaged, mixed frequency and phase recognition accuracy across the subjects and (b) that of each individual subject with respect to the data length using the proposed decoder, respectively. The legend of (b) is consistent with the one in (a).

Chapter 7

Conclusions and Open Problems

7.1 Conclusions

We conclude the study in this section. In this thesis, we have discussed unexplored problems and challenges on SSVEP-based BMI. The methods have proposed to address the problems;

1) comparison with ETI

2) calibration reduction

3) achieving higher recognition accuracy

4) asynchronousness

In Chapter 3, we compared between SSVEP-based BMI and dwelling-based ETI, and clarified their drawbacks and advantages. Their performances were evaluated by investigating their accuracies and information transfer rate (ITR) with respect to the target size and the command analysis time, i.e., time window length of EEG analysis or the dwell time. Recognition accuracy and the ITR were measured by giving subjects the task of selecting one of four targets by gazing at it. The targets were displayed in three different sizes, with sides 20, 40, and 60 mm long, to evaluate performance with respect to the target size. The experimental results showed that the BMI was comparable to the ETI in terms of accuracy and the information transfer rate.

In particular, when the size of a target was relatively small, the BMI had significantly better performance than the ETI.

In Chapter 4, the method to reduce calibration time for SSVEP-based BMI has been proposed. The proposed method attempts to generate from source templates to target templates corresponding to the rest of commands. The target templates were obtained by shifting the frequency and phase of the source template to the desired frequency and phase. The experimental results suggested that the proposed transferred template enabled to recognize both frequency and phase with a small amount of calibration closely achieving the performance using the full calibration data.

In Chapter 5, in order to detect commands fast and accurately so as to implement mixed-coded SSVEP-BMI as a reliable interface, we presented a novel method which employed multiset canonical correlation analysis (MCCA) to obtain spatial filters which enhance SSVEP components. An experiment with a mixed-coded SSVEP-BMI was conducted to evaluate performance of the proposed method compared with the previous work. The experimental results showed a significant improvement in command recognition accuracy and ITR with the proposed method compared to the state-of-the-art. These results indicate the efficacy of employing MCCA to recognize the frequency and the phase of SSVEPs.

In Chapter 6, we proposed a novel asynchronous BMI implemented with mixed frequency and phase-coded visual stimuli. The mixed-coded visual stimuli were presented as flickers with a following blank interval to synchronize the recorder of EEG with the stimuli, which was aimed to detect the phase in an asynchronous situation. For decoding from the measured EEG, MCCA was efficiently exploited for recognizing the intentional state and the intending command. The proposed asynchronous BMI was tested on 11 healthy subjects. The proposed decoder was capable of discriminating

between the IC/NC state and determining the command faster and more accurately than the contrast methods. The BMI based on mixed-coded visual stimuli was able to be implemented in an asynchronous manner, and the MCCA-based decoder outperformed the conventional ones in terms of discriminability of intentional states and command recognition accuracy. The present study showed that an asynchronous BMI can be implemented with mixed-coded visual stimuli for the first time, which enables a large increase in the number of choices/commands.

7.2 Open Problems

Finally, we present some open problems on SSVEP-based BMIs. Although we have addressed the solutions for several problems, the following issues are remaining.

7.2.1 Utilizing higher stimulus frequencies

In SSVEP-based BMI, flickering stimuli can cause visual fatigue. Additionally, there is a risk of photosensitive epileptic seizures for a cirtain people. Therefore, using higher frequency for visual stimuli have adavantages in terms of reducing visual fatigue and the risk of epileptic seizures. Stimulus frequencies to elicit SSVEP can be classified into three tanges: low (< 12 Hz), medium (12–30 Hz), and high (> 30 Hz) [20]. Generally, SSVEPs elicited by high stimulus frequencies have smaller amplitide and more difficut to detect compared to low and medium frequency ranges. There have been several studies reported to tackle this challenge [119, 120, 56]. For example, Sakurada *et al.*reported that frequenct recogntion accuracy achieved 93.1% using 41, 43, and 45 Hz of stimulus frequencies [56]. Although it showed an inportant possibility using high stimulus frequencies in SSVEP-based BMI, it also showed the difficulties to employ those frequencies, such as, frequency

recognition accuracy was lower (88.0%) with invisible stimulus frequencies, 61, 63, and 65 Hz; the number of classes is small; a relatively long EEG signal (3–5 seconds) is required for analysis, which leads to a longer time to input a command.

7.2.2 Employing training data of other users for reference signals

As described in Chapter 4, 5, and 6, utilizing individual templates as reference signals is a promising method to achieve high performance in SSVEP-based BMI. To generate individual template, we need to collect calbration dataset from each individual user, which can take a lot of time. In Chapter 4, we proposed a method to reduce calibration by generating new templates from small calibration dataset. On the other hand, some researchers have attempted to reduce calibration by utilizing other subjects' calibration data to artificially generate reference signals [91, 121]. For example, Yuan *et al.*generated new templates by averaging training EEG signals of exiting subjects (source subjects) across multiple subjects and trials [91]. The recognition accuracy was enhanced by 20.94% compared to the standard CCA method [39] where the number of the source subjects was 11. EEG signals of each source subject were collected from 240 trials. These researches showed a potential to reduce calibration and still achieve a good performance by utilizing other users' training data, although, it is a challenge how to collect enough training data as an open dataset of the EEG signals (especially SSVEP) still remails few.

Bibliography

[1] Jean-Jacques Vidal. "Toward direct brain–computer communication". *Annu. Review of Biophysics and Bioeng.* 2.1 (1973), pp. 157–180.

[2] J R Wolpaw et al. "Brain-computer interfaces for communication and control". *Clin. Neurophysiol.* 113.6 (June 2002), pp. 767–791. ISSN: 1388-2457.

[3] Melody M Moore. "Real-world applications for brain-computer interface technology". *IEEE Trans. Neural Syst. Rehabil. Eng.* 11.2 (2003), pp. 162–165.

[4] A Kübler et al. "BCI meeting 2005—workshop on clinical issues and applications". *IEEE Trans. Neural Syst. Rehabil. Eng.* 14.2 (2006), pp. 131–134.

[5] Mikhail A Lebedev and Miguel AL Nicolelis. "Brain–machine interfaces: Past, present and future". *TRENDS in Neurosciences* 29.9 (2006), pp. 536–546.

[6] Ali Bashashati et al. "A survey of signal processing algorithms in brain–computer interfaces based on electrical brain signals". *Journal of Neural engineering* 4.2 (2007), R32.

[7] Andrzej Cichocki et al. "Noninvasive BCIs: Multiway signal-processing array decompositions". *Computer* 10 (2008), pp. 34–42.

[8] Janis J Daly and Jonathan R Wolpaw. "Brain–computer interfaces in neurological rehabilitation". *The Lancet Neurology* 7.11 (2008), pp. 1032–1043.

[9] Wei Wang et al. “Neural interface technology for rehabilitation: Exploiting and promoting neuroplasticity”. *Physical Medicine and Rehabilitation Clinics of North America* 21.1 (2010), pp. 157–178.

[10] Thorsten O Zander and Christian Kothe. “Towards passive brain–computer interfaces: Applying brain–computer interface technology to human–machine systems in general”. *Journal of Neural Engineering* 8.2 (2011), p. 025005.

[11] Dennis J McFarland and Jonathan R Wolpaw. “Brain–computer interfaces for communication and control”. *Communications of the ACM* 54.5 (2011), pp. 60–66.

[12] L.A. Farwell and E. Donchin. “Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials”. *Electroencephalography and Clinical Neurophysiology* 70.6 (Dec. 1988), pp. 510–523.

[13] Benjamin Blankertz et al. “Single-trial analysis and classification of ERP componentsa tutorial”. *NeuroImage* 56.2 (2011), pp. 814–825.

[14] Andrea Kübler et al. “A brain–computer interface controlled auditory event-related potential (P300) spelling system for locked-in patients”. *Annals of the New York Academy of Sciences* 1157.1 (2009), pp. 90–100.

[15] Marjolein van der Waal et al. “Introducing the tactile speller: An ERP-based brain–computer interface for communication”. *Journal of Neural Engineering* 9.4 (2012), p. 045002.

[16] Erich E Sutter. “The brain response interface: Communication through visually-induced electrical brain responses”. *Journal of Microcomputer Applications* 15.1 (1992), pp. 31–45.

[17] Robert F Burkard, Jos J Eggermont, and Manuel Don. *Auditory evoked potentials: Basic principles and clinical application*. Lippincott Williams & Wilkins, 2007.

[18] Matthew Middendorf et al. "Brain–computer interfaces based on the steady-state visual-evoked response". 8.2 (2000), pp. 211–214.

[19] François-Benoît Vialatte et al. "Steady-state visually evoked potentials: Focus on essential paradigms and future perspectives". *Progress in Neurobiology* 90.4 (2010), pp. 418–438.

[20] David Regan. *Human brain electrophysiology: Evoked potentials and evoked magnetic fields in science and medicine*. Elsevier, 1989.

[21] Christoph S Herrmann. "Human EEG responses to 1–100 Hz flicker: Resonance phenomena in visual cortex and their potential correlation to cognitive phenomena". *Experimental Brain Research* 137.3-4 (2001), pp. 346–353.

[22] Shravani Sur and VK Sinha. "Event-related potential: An overview". *Industrial psychiatry journal* 18.1 (2009), p. 70.

[23] G Pfurtscheller et al. "EEG-based discrimination between imagination of right and left hand movement". *Electroencephalography and clinical Neurophysiology* 103.6 (1997), pp. 642–651.

[24] Saeid Sanei. *Adaptive processing of brain signals*. John Wiley & Sons, 2013.

[25] Shangkai Gao et al. "Visual and auditory brain–computer interfaces". *IEEE Trans. Biomed. Eng.* 61.5 (Jan. 2014), pp. 1436–1447.

[26] Xiaogang Chen et al. "High-speed spelling with a noninvasive brain–computer interface". *Proceedings of the National Academy of Sciences* 112.44 (2015), E6058–E6067.

[27] Arie A Kaufman, Amit Bandopadhay, and George J Piligian. *Apparatus and method for eye tracking interface*. 1994.

[28] OK Oyekoya and FWM Stentiford. "Eye tracking as a new interface for image retrieval". *BT Tech. J.* 22.3 (July 2004), pp. 161–169.

[29] Oyewole Oyekoya and Fred Stentiford. "An eye tracking interface for image search". In: *Proc. Symp. Eye Tracking Research & Applicat.* Mar. 2006, pp. 40–40.

[30] Weston Sewell and Oleg Komogortsev. "Real-time eye gaze tracking with an unmodified commodity webcam employing a neural network". In: *CHI'10 Extended Abstracts Human Factors Comput. Syst.* Apr. 2010, pp. 3739–3744.

[31] Michael Ashmore, Andrew T Duchowski, and Garth Shoemaker. "Efficient eye pointing with a fisheye lens". In: *Proc. Graph. Interface 2005*. Canadian Human-Computer Communications Society. 2005, pp. 203–210.

[32] R Spataro et al. "The eye-tracking computer device for communication in amyotrophic lateral sclerosis". *Acta Neurol. Scand.* 130.1 (2014), pp. 40–45.

[33] Brendan Allison et al. "BCI demographics: How many (and what kinds of) people can use an SSVEP BCI?" *IEEE Trans. Neural Syst. and Rehabil. Eng.* 18.2 (2010), pp. 107–116.

[34] M Middendorf et al. "Brain-computer interfaces based on the steady-state visual-evoked response". *IEEE Trans. Rehabil. Eng.* 8.2 (June 2000), pp. 211–4.

[35] Colin Ware and Harutune H. Mikaelian. "An evaluation of an eye tracker as a device for computer input". In: *Proc. SGICHI/GI Conf. Human Factors Computing Syst. and Graph. Interface*. May 1987, pp. 183–188.

[36] Kaori Suefusa and Toshihisa Tanaka. "Reduced calibration by efficient transformation of templates for high speed hybrid coded SSVEP brain-computer interfaces". In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2017*. IEEE. 2017, pp. 929–933.

[37] Masaki Nakanishi et al. "A high-speed brain speller using steady-state visual evoked potentials". *International Journal of Neural Systems* 24.06 (2014), p. 1450019.

[38] Harold Hotelling. "Relations between two sets of variates". *Biometrika* 28.3 (Dec. 1936), pp. 321–377.

[39] Zhonglin Lin et al. "Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs". *IEEE Trans. Biomed. Eng.* 53.12 (Dec. 2006), pp. 2610–2614.

[40] José del R Millán and Josep Mouriño. "Asynchronous BCI and local neural classifiers: An overview of the adaptive brain interface project". *IEEE Trans. Neural Syst. Rehab. Eng.* 11.2 (June 2003), pp. 159–161.

[41] Hubert Cecotti. "A self-paced and calibration-less SSVEP-based brain–computer interface speller". *IEEE Trans. Neural Syst. Rehab. Eng.* 18.2 (Apr. 2010), pp. 127–133.

[42] Aneesa Akmal Nooh, Jasmy Yunus, and Salwani Mohd Daud. "A review of asynchronous electroencephalogram-based brain computer interface systems". In: *Int. Conf. Biomed. Eng. Tech. IPCBEE*. Vol. 11. 2011, pp. 55–59.

[43] Ming Cheng et al. “Design and implementation of a brain–computer interface with high transfer rates”. *IEEE Transactions on Biomedical Engineering* 49.10 (2002), pp. 1181–1186.

[44] Sergio Parini et al. “A robust and self-paced BCI system based on a four class SSVEP paradigm: Algorithms and protocols for a high-transfer-rate direct brain communication”. *Computational Intell. Neuroscience* 2009 (Feb. 2009).

[45] Gert Pfurtscheller et al. “The hybrid BCI”. *Frontiers in Neuroscience* 4 (2010).

[46] Bin Xia et al. “Asynchronous brain–computer interface based on steady-state visual-evoked potential”. *Cognitive Comput.* 5.2 (June 2013), pp. 243–251.

[47] Jon R Kettenring. “Canonical analysis of several sets of variables”. *Biometrika* 58.3 (1971), pp. 433–451.

[48] Chuan Jia et al. “Frequency and phase mixed coding in SSVEP-based brain–computer interface”. *IEEE Trans. Biomed. Eng.* 58.1 (2011), pp. 200–206.

[49] Kaori Suefusa and Toshihisa Tanaka. “Decoding of responses to mixed frequency and phase coded visual stimuli using multiset canonical correlation analysis”. In: *2016 38th Annu. Int. Conf. IEEE Eng. Med. Biology Soc.* IEEE. 2016, pp. 1492–1495.

[50] *ALS INSTITUTE*. https://www.als.net/about-als-tdi/als-faq/. Accessed: 2019-02-04.

[51] Alois Schlögl, Mel Slater, and Gert Pfurtscheller. “Presence research and EEG”. In: *Proceedings of the 5th International Workshop on Presence*. 2002, pp. 9–11.

[52] Melody Moore Jackson and Rudolph Mappus. "Neural control interfaces". In: *Brain-Computer Interfaces*. Springer, 2010, pp. 21–33.

[53] Reinhold Scherer et al. "The self-paced Graz brain-computer interface: Methods and applications". *Computational Intell. Neuroscience* 2007 (July 2007).

[54] Robert Leeb et al. "Self-paced (asynchronous) BCI control of a wheelchair in virtual environments: A case study with a tetraplegic". *Computational Intell. Neuroscience* 2007 (July 2007).

[55] Danhua Zhu et al. "A survey of stimulation methods used in SSVEP-based BCIs". *Intell. Neurosci.* 2010 (Jan. 2010), pp. 1–12.

[56] Takeshi Sakurada et al. "Use of high-frequency visual stimuli above the critical flicker frequency in a SSVEP-based BMI". *Clin. Neurophysiol.* 126.10 (2015), pp. 1972–1978.

[57] Abdul Moiz Penkar, Christof Lutteroth, and Gerald Weber. "Designing for the eye - design parameters for dwell in gaze interaction". In: *Proc. 24th Australian Comput.-Human Interaction Conf.* Nov. 2012, pp. 479–488.

[58] Aulikki Hyrskykari, Howell Istance, and Stephen Vickers. "Gaze gestures or dwell-based interaction?" In: *Proc. Symp. Eye Tracking Research & Applicat.* Santa Barbara, California, Mar. 2012, pp. 229–232. ISBN: 9781450312257.

[59] Shumin Zhai, Carlos Morimoto, and Steven Ihde. "Manual and Gaze Input Cascaded (MAGIC) Pointing". In: *Proc. SIGCHI Conf. Human Factors Computing Syst.* CHI '99. Pittsburgh, Pennsylvania, USA, May 1999, pp. 246–253. ISBN: 0-201-48559-1.

[60] Morten Lund Dybdal, Javier San Agustin, and John Paulin Hansen. "Gaze input for mobile devices by dwell and gestures". In: *Proc. Symp.*

Eye Tracking Research & Applicat. Santa Barbara, California, Mar. 2012, pp. 225–228. ISBN: 978-1-4503-1221-9.

[61] Xiaogang Chen et al. "High-speed spelling with a noninvasive brain–computer interface". *Proc. National Acad. Sci.* 112.44 (2015), E6058–E6067.

[62] Päivi Majaranta and Kari-Jouko Räihä. "Text entry by gaze: Utilizing eye-tracking". *Text Entry Systems: Mobility, Accessibility, Universality* (2007), pp. 175–187.

[63] Ivo Käthner, Andrea Kübler, and Sebastian Halder. "Comparison of eye tracking, electrooculography and an auditory brain-computer interface for binary communication: A case study with a participant in the locked-in state". *J. Neuroeng. and Rehabil.* 12.1 (2015), p. 1.

[64] Emanuele Pasqualotto et al. "Usability and workload of access technology for people with severe motor impairment: A comparison of brain-computer interfacing and eye tracking". *Neurorehabil. and Neural Repair* (2015), p. 1545968315575611.

[65] J.R. Wolpaw et al. "EEG-based communication: Improved accuracy by response verification". *IEEE Trans. Rehabil. Eng.* 6.3 (1998), pp. 326–333.

[66] Kaori Suefusa and Toshihisa Tanaka. "Visually stimulated brain-computer interfaces compete with eye tracking interfaces when using small targets". In: *2014 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*. IEEE. 2014, pp. 4005–4008.

[67] *Scene Camera Eye Tracking System Specifications — Arrington Research*. http://www.arringtonresearch.com/scenetechinfo.html. Accessed: 2016-10-1.

[68] Dongheng Li, Jason Babcock, and Derrick J. Parkhurst. "openEyes: A low-cost head-mounted eye-tracking solution". In: *Proc. Symp. Eye Tracking Research & Applicat.* San Diego, California, Mar. 2006, pp. 95–100.

[69] Hovagim Bakardjian, Toshihisa Tanaka, and Andrzej Cichocki. "Optimization of SSVEP brain responses with application to eight-command Brain–Computer Interface". *Neurosci. Lett.* 469.1 (Jan. 2010), pp. 34–38.

[70] Xiaogang Chen et al. "Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain–computer interface". *J. Neural Eng.* 12.4 (June 2015), p. 046008.

[71] Guangyu Bin et al. "A high-speed BCI based on code modulation VEP". *J. Neural Eng.* 8.2 (Mar. 2011), p. 025015.

[72] Yu Zhang et al. "Multiway canonical correlation analysis for frequency components recognition in SSVEP-based BCIs". In: *Int. Conf. on Neural Informat. Process.* Springer. 2011, pp. 287–295.

[73] Yu Zhang et al. "L1-regularized multiway canonical correlation analysis for SSVEP-based BCI". *IEEE Trans. Neural Syst. and Rehabil. Eng.* 21.6 (2013), pp. 887–896.

[74] Yu Zhang et al. "Frequency recognition in SSVEP-based BCI using multiset canonical correlation analysis". *Int. J. Neural Syst.* 24.04 (2014), p. 1450013.

[75] Xiaogang Chen et al. "Hybrid frequency and phase coding for a high-speed SSVEP-based BCI speller". In: *2014 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*. IEEE. 2014, pp. 3993–3996.

[76] Zhonglin Lin et al. "Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs". *IEEE Trans. Biomed. Eng.* 53.12 (Dec. 2006), pp. 2610–2614.

[77] Chris Lankford. "Effective eye-gaze input into windows". In: *Proc. 2000 Symp. Eye Tracking Research & Applicat.* Palm Beach Gardens, Florida, USA: ACM, Nov. 2000, pp. 23–27.

[78] John Paulin Hansen et al. "Command without a click: Dwell time typing by mouse and gaze selections". In: *Proc. Human–Comput. Interaction – INTERACT' 03*. Sept. 2003, pp. 121–128.

[79] Y. Kimura et al. "SSVEP-based brain–computer interfaces using FSK-modulated visual stimuli". *IEEE Trans. Biomed. Eng.* 60.10 (Oct. 2013), pp. 2831–2838.

[80] Masaki Nakanishi et al. "A high-speed brain speller using steady-state visual evoked potentials". *Int. J. Neural Syst.* 24.6 (Sept. 2014), p. 1450019.

[81] Robert J. K. Jacob. "The use of eye movements in human-computer interaction techniques: What you look at is what you get". *ACM Trans. Inf. Syst.* 9.2 (Apr. 1991), pp. 152–169. ISSN: 1046-8188.

[82] *ViewConfiguration — Android Developers*. http://developer.android.com/reference/android/view/ViewConfiguration.html. Accessed: 2016-4-30.

[83] G. Bartl, G. H. Van Lith, and G. W. Van Marle. "Cortical potentials evoked by a TV pattern reversal stimulus with varying check sizes and stimulus field". *British J. Ophthalmology* 62.4 (Apr. 1978), pp. 216–219.

[84] Anna Duszyk et al. "Towards an optimization of stimulus parameters for brain-computer interfaces based on steady state visual evoked potentials". *PLoS ONE* 9.11 (Nov. 2014), p. 112099.

[85] Minho Kim, Byung Hyung Kim, and Sungho Jo. "Quantitative evaluation of a low-cost noninvasive hybrid interface based on EEG and

eye movement". *IEEE Trans. Neural Syst. and Rehabil. Eng.* 23.2 (Mar. 2015), pp. 159–168.

[86] Päivi Majaranta and Kari-Jouko Räihä. "Twenty years of eye typing: Systems and design issues". In: *Proc. Symp. Eye Tracking Research & Applicat.* New Orleans, Louisiana, Mar. 2002, pp. 15–22. ISBN: 1-58113-467-3.

[87] Xiaogang Chen et al. "Hybrid frequency and phase coding for a high-speed SSVEP-based BCI speller". In: *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE. 2014, pp. 3993–3996.

[88] Chuan Jia et al. "Frequency and phase mixed coding in SSVEP-based brain–computer interface". *IEEE Transactions on Biomedical Engineering* 58.1 (2011), pp. 200–206.

[89] Yu-Te Wang et al. "Developing an online steady-state visual evoked potential-based brain-computer interface system using EarEEG". In: *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE. 2015, pp. 2271–2274.

[90] Kaori Suefusa and Toshihisa Tanaka. "Decoding of responses to mixed frequency and phase coded visual stimuli using multiset canonical correlation analysis". In: *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE. 2016, pp. 1492–1495.

[91] Peng Yuan et al. "Enhancing performances of SSVEP-based brain–computer interfaces via exploiting inter-subject information". *Journal of Neural Engineering* 12.4 (2015), p. 046006.

[92] Yijun Wang, Y-T Wang, and T-P Jung. "Visual stimulus design for high-rate SSVEP BCI". *Electronics letters* 46.15 (2010), pp. 1057–1058.

[93] Robert Oostenveld and Peter Praamstra. "The five percent electrode system for high-resolution EEG and ERP measurements". *Clinical Neurophysiology* 112.4 (2001), pp. 713–719.

[94] Javier Vía, Ignacio Santamaría, and Jesús Pérez. "A learning algorithm for adaptive canonical correlation analysis of several data sets". *Neural Networks* 20.1 (2007), pp. 139–152.

[95] Allan Aasbjerg Nielsen. "Multiset canonical correlations analysis and multispectral, truly multitemporal remote sensing data". *IEEE Transactions on Image Processing* 11.3 (2002), pp. 293–305.

[96] YU Zhang et al. "Frequency recognition in SSVEP-based BCI using multiset canonical correlation analysis". *International Journal of Neural Systems* 24.04 (2014), p. 1450013.

[97] Yi-Ou Li et al. "Joint blind source separation by multiset canonical correlation analysis". *IEEE Transactions on Signal Processing* 57.10 (2009), pp. 3918–3929.

[98] Masaki Nakanishi et al. "A dynamic stopping method for improving performance of steady-state visual evoked potential based brain-computer interfaces". In: *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE. 2015, pp. 1057–1060.

[99] Hovagim Bakardjian, Toshihisa Tanaka, and Andrzej Cichocki. "Optimization of SSVEP brain responses with application to eight-command brain–computer interface". *Neuroscience Letters* 469.1 (2010), pp. 34–38.

[100] Pawel Poryzala and Andrzej Materka. "Cluster analysis of CCA coefficients for robust detection of the asynchronous SSVEPs in brain–computer interfaces". *Biomedical Signal Process. and Control* 10 (2014), pp. 201–208.

[101] Zoltan Joseph Koles. "The quantitative extraction and topographic mapping of the abnormal components in the clinical EEG". *Electroencephalography and Clinical Neurophysiology* 79.6 (1991), pp. 440–447.

[102] Ola Friman, Ivan Volosyak, and Axel Gräser. "Multiple channel detection of steady-state visual evoked potentials for brain–computer interfaces". *IEEE Transactions on Biomedical Engineering* 54.4 (2007), pp. 742–750.

[103] GE Chatrian, E Lettich, and PL Nelson. "Ten percent electrode system for topographic studies of spontaneous and evoked EEG activities". *American Journal of EEG technology* 25.2 (1985), pp. 83–92.

[104] Christopher M Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

[105] Yaqian Guo, Trevor Hastie, and Robert Tibshirani. "Regularized discriminant analysis and its application in microarrays". *Biostatistics* 1.1 (2005), pp. 1–18.

[106] Yoav Freund and Robert E Schapire. "A decision-theoretic generalization of on-line learning and an application to boosting". In: *European Conf. Computational Learning Theory*. Springer. 1995, pp. 23–37.

[107] Milton Friedman. "The use of ranks to avoid the assumption of normality implicit in the analysis of variance". *J. Amer. Statistical Assoc.* 32.200 (1937), pp. 675–701.

[108] Sidney Siegel and Castellan N. John Jr. *Nonparametric statistics for the behavioral sciences*. McGraw-hill, 1988.

[109] Andy Field. *Discovering statistics using SPSS*. Sage Publications, 2009.

[110] A Kübler et al. "A model of BCI-control". In: *Proc. 5th Int. Brain–Computer Interface Conf.* 2011, pp. 100–103.

[111] F Aloise et al. "P300-based brain–computer interface for environmental control: An asynchronous approach". *J. Neural Eng.* 8.2 (2011), p. 025025.

[112] Karl LaFleur et al. "Quadcopter control in three-dimensional space using a noninvasive motor imagery-based brain–computer interface". *J. Neural Eng.* 10.4 (2013), p. 046003.

[113] Bonkon Koo et al. "A hybrid NIRS-EEG system for self-paced brain computer interface with online motor imagery". *J. Neuroscience Methods* 244 (2015), pp. 26–32.

[114] Luzheng Bi, Xin-An Fan, and Yili Liu. "EEG-based brain-controlled mobile robots: A survey". *IEEE Trans. Human-Machine Syst.* 43.2 (2013), pp. 161–176.

[115] Haihong Zhang, Cuntai Guan, and Chuanchu Wang. "Asynchronous P300-based brain–computer interfaces: A computational approach with statistical models". *IEEE Trans. Biomed. Eng.* 55.6 (2008), pp. 1754–1763.

[116] Rajesh C Panicker, Sadasivan Puthusserypady, and Ying Sun. "An asynchronous P300 BCI with SSVEP-based control state detection". *IEEE Trans. Biomed. Eng.* 58.6 (2011), pp. 1781–1788.

[117] Yuanqing Li et al. "A hybrid BCI system combining P300 and SSVEP and its application to wheelchair control". *IEEE Trans. Biomed. Eng.* 60.11 (2013), pp. 3156–3166.

[118] Hannes Riechmann et al. "Asynchronous, parallel on-line classification of P300 and ERD for an efficient hybrid BCI". In: *2011 5th Int. IEEE/EMBS Conf. Neural Eng. (NER)*. IEEE. 2011, pp. 412–415.

[119] Pablo F Diez et al. "Asynchronous BCI control using high-frequency SSVEP". *Journal of Neuroengineering and Rehabilitation* 8.1 (2011), p. 39.

[120] Ivan Volosyak et al. "BCI demographics II: How many (and what kinds of) people can use a high-frequency SSVEP BCI?" *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 19.3 (2011), pp. 232–239.

[121] Nicholas R Waytowich et al. "Spectral transfer learning using information geometry for a user-independent brain-computer interface". *Frontiers in Neuroscience* 10 (2016), p. 430.

Appendix A

List of Publications

Journal Papers

1. **Kaori Suefusa** and Toshihisa Tanaka, "Asynchronous brain–computer interfacing based on mixed-coded visual stimuli," IEEE Transactions on Biomedical Engineering, vol. 65, no. 9, pp. 2119–2129, 2018. (Chapter 6).

2. Haruki Okawa, **Kaori Suefusa**, and Toshihisa Tanaka, "Neural entrainment to auditory imagery of rhythms," Frontiers in Human Neuroscience, vol. 11, p. 493, 2017.

3. **Kaori Suefusa** and Toshihisa Tanaka, "A comparison study of visually stimulated brain–computer and eye-tracking interfaces," Journal of Neural Engineering, vol. 14, no. 3. p. 036009, 2017. (Chapter 3)

International Conference Papers

1. Yutaro Tanji, Masaki Nakanishi, **Kaori Suefusa**, and Toshihisa Tanaka, "Waveform-based multi-stimulus coding for braincomputer interfaces based on steady-state visual evoked potentials," in Proc. of the 43rd IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2018), Calgary, Canada, Apr. 2018.

2. **Kaori Suefusa** and Toshihisa Tanaka, "Reduced calibration by efficient transformation of templates for high speed hybrid coded SSVEP brain–computer interfaces," in Proc. of the 42nd IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2017), New Orleans, USA, Mar. 2017, pp. 929–933. (Chapter 4)

3. **Kaori Suefusa** and Toshihisa Tanaka, "Decoding of responses to mixed frequency and phase coded visual stimuli using multiset canonical correlation analysis," in Proc. of 38th Annual Conference of IEEE Engineering in Medicine and Biology Society (EMBC 2016), Orland, USA, Aug. 2016, pp. 1492–1495. (Chapter 5)

4. **Kaori Suefusa** and Toshihisa Tanaka, "Phase-based detection of intentional state for asynchronous brain–computer interface," in Proc. of the 40th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2015), Brisbane, Australia, Apr. 2015, pp. 808–812.

5. **Kaori Suefusa** and Toshihisa Tanaka, "Visually stimulated brain–computer interfaces compete with eye tracking interfaces when using small targets," in Proc. of 36th Annual Conference of IEEE Engineering in Medicine and Biology Society (EMBC 2014), Chicago, USA, Aug. 2014, pp. 4005–4008.

Poster Presentations

1. **Kaori Suefusa** and Toshihisa Tanaka, "Calibration reduction by efficient transformation of individual templates for SSVEP Brain–Computer Interface,"

in Proc. of the 10th Asia-Pacific Signal and Information Processing Association Annual Summit and Conference 2018, Hawaii, USA, Oct. 2018.

2. Yutaro Tanji, Naoki Morikawa, Masaki Nakanishi, **Kaori Suefusa**, and Toshihisa Tanaka, "Classifying modulation waveforms of visual stimuli via steady-state visual evoked potentials," in Proc. of 38th Annual Conference of IEEE Engineering in Medicine and Biology Society (EMBC 2017), Jeju Island, South Korea, Jul. 2017.

3. Haruki Okawa, **Kaori Suefusa**, and Toshihisa Tanaka, "Imagined rhythm can be decoded from EEG," Neuroscience 2016, San Diego, USA, Nov. 2016.

4. **Kaori Suefusa** and Toshihisa Tanaka, "Asynchronous brain–computer interface using canonical correlation analysis," in Proc. of the Organization for Human Brain Mapping (OHBM) 2015, Honolulu, USA, Jun. 2015, p. 2371.

Domestic Conference Papers

1. **末房佳小里**，田中聡久，"視覚刺激を用いた脳コンピュータインタフェースにおけるキャリブレーション削減のための参照信号の効果的な変換法，" 電子情報通信学会第 31 回信号処理シンポジウム論文集，pp. 67–68，吹田市，2016 年 11 月.

2. 大川陽貴，**末房佳小里**，田中聡久，"脳波からのリズム想像デコーディング" 電気情報通信学会技術研究報告，vol. 116, no. 95, SIP2016-64, pp. 161–166，弘前市，2016 年 6 月.

3. **末房佳小里**，田中聡久，“マルチセット正準相関分析による定常的視覚誘発電位の周波数と位相の識別，” 第 55 回日本生体医工学会大会抄録集，富山市，2016 年 4 月．

4. **末房佳小里**，田中聡久，“マルチセット正準相関分析による定常的視覚誘発電位の周波数識別，” 電子情報通信学会第 30 回信号処理シンポジウム論文集，pp. 68–69，いわき市，2015 年 11 月．

5. **末房佳小里**，田中聡久，“位相空間における正準相関分析を利用した非同期脳コンピュータインタフェース，” 電子情報通信学会技術研究報告，vol. 114, no. 473, SIP2014-141, pp. 151–156，石垣市，2015 年 3 月．

6. **末房佳小里**，新保圭太，田中聡久，“非同期脳コンピュータインタフェースにおけるユーザ依存のパラメータ選択法，” 電子情報通信学会技術研究報告，vol. 114, no. 122, SIP2014-55, pp. 177–181, 札幌市，2014 年 7 月．

7. **末房佳小里**，田中聡久，“コマンド入力インタフェースにおける視線位置と脳波利用の性能評価，” 電子情報通信学会第 28 回信号処理シンポジウム論文集，pp. 100–105, 下関市，2013 年 11 月．