TOWARDS PREDICTION AND IMPROVEMENT OF EEG-BASED MI-BCI PERFORMANCE

Atieh Bamdadian

(B.Sc. and M.Sc. (Biomedical Engineering))

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2014

Declaration

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Atieh Bamdadian

13.02.2015

Acknowledgments

I would like to express my very deep gratitude to my supervisor Dr. Cuntai Guan first of all because of all his support and guidance over the past four years. I'm really grateful of having an opportunity to do research under his supervision. I would also like to deeply thank my supervisor Prof. Jianxin Xu, for his constructive guidance during my PhD studies. Many thanks also go to Dr. Kai Keng Ang because of all his kindness and unlimited assistance and support. I still have a lot to learn from him.

I would like to thank my friends and colleagues in Brain-Computer Interface (BCI) Laboratory at Institute for Infocomm Research (I2R). It was an honor for me to be a member of such a scientific and friendly research group. Many thanks also go to my lab-mates for being compassionate and supportive especially during the last few months that I was conducting my experiment.

I would also like to express my warmest gratitude to A*STAR for offering me SINGA scholarship and fund my research.

Last but not the least I would like to thank my parents for their unconditional support during my life, thanks for always encouraging and helping me achieve my dreams.

Contents

Summary

A brain-computer interface (BCI) is a communication device controlled by brain activity without any actual muscle movement. Motor imagery (MI) is one of the mental activities to control a BCI which induces changes in sensorimotor rhythms (SMRs) of electroencephalogram (EEG). A MI-BCI performance reflects how well a subject can control an EEG-based MI-BCI. However, there is a large performance variation among BCI users, and the reason why some subjects cannot use MI-BCI to achieve even moderate performance is still unknown. One of the possible reasons, known as BCI deficiency in the literature, is the subjects' inability in modulating their brain rhythms. Another reason is the non-stationarity of EEG signal.

The main goal of this thesis is to improve the MI-BCI performance of the subjects. To do so, we can help subjects with BCI deficiency improve their performance, and develop adaptive algorithms. In this thesis, novel neurophysiological predictors are proposed that are correlated to the subject's BCI performance, and neurofeedback training using these predictors is proposed to enhance the MI-BCI performance of subjects. Subsequently, two adaptive algorithms are proposed to address inter-session non-stationarity and to enhance the performance of the subjects.

The first predictor is a novel neurophysiological coefficient computed from the spectral power of pre-cue EEG data over different regions of the brain. The proposed neurophysiological coefficient is a possible quantification of the attention level of the subject that captures spatial information from the brain in addition to spectral information of alpha, beta, and theta frequency bands. The second proposed neurophysiological predictor uses spatio-spectral decomposition (SSD) to extract the spatial-spectral compo-

nents from pre-cue and resting state EEG. The extracted spatial-spectral components have higher signal-to-noise ratio to improve the performance prediction of subjects using MI-BCI. A new experiment is also designed to help subjects improve their BCI performance through enhancing their resting state alpha SSD components in multiple neurofeedback training sessions.

Two adaptive algorithms are also developed to address inter-session non-stationarity. The first proposed algorithm is based on Kullback-Leibler (KL)-distance weighting to adapt features extracted using filter bank common spatial pattern (FBCSP). The second proposed algorithm, adaptive extreme learning machine (A-ELM), is a method that adapts the initial classifier to overcome the features' drift from calibration session using limited number of data from the evaluation session.

The proposed algorithms were applied on the data collected from healthy subjects. The results showed that the proposed predictors successfully assessed the MI-BCI performance. Moreover, the results revealed that low attention level and low alpha band power in resting state resulted in poor performance. In addition, the neurofeedback training enhanced the resting sate alpha activity that subsequently led to improved MI-BCI performance. Finally, by adapting the features and classifier, the proposed algorithms KL-distance weighting of FBCSP and A-ELM, also improved the accuracy of the EEG-based MI-BCI.

In conclusion, the proposed predictors successfully detect BCI deficient subjects. The results showed that the MI-BCI performance of subjects can be improved from neurofeedback training using the proposed alpha SSD-based predictor and by using the proposed adaptive algorithms. The proposed algorithms can be applied in real world applications to find out whether MI-BCI is a suitable therapy method for disabled people, and also to help subjects with BCI deficiency to better control a MI-BCI system.

List of Tables

List of Figures

Acronyms

2HAND Two-Hand Coordination Test

A-ELM Adaptive Extreme Learning Machine

ADHD Attention Defcit Hyperactivity Disorder

AHA Attitude Towards Work

ALS Amyotrophic Lateral Sclerosis

AMT Active Movement Training

ANOVA Analysis of Variance

BCI Brain-Computer Interface

BOLD Blood Oxygen Level Dependent

CAR Common Average Reference

CCSP Composite Common Spatial Pattern

CSP Common Spatial Pattern

CV Cross-Validation

DSA Domain Space Adaptation

ECoG Electrocorticography

EEG Electroencephalography

ELM Extreme Learning Machine

EPs Evoked Potentials

FBCSP Filter Bank Common Spatial Pattern

fMRI Functional Magnetic Resonance Imaging

fMRI Functional Near-Infrared Spectroscopy

ICA Independent Component Analysis

KL Kullback-Leibler

LAR Local Average Reference

LDA Linear Discriminant Analysis

ME Motor Execution

MEG Magnetoencephalography

MI Motor Imagery

NBPW Naïve Bayesian Parzen Window

NFT Neurofeedback training

PCA Principal Component Analysis

PM Passive Movement

SCPs Slow Cortical Potentials

SL Surface Laplacian

SLFN Single-Hidden Layer Feed-Forward Neural Network

SMRs Sensorimotor Rhythms

SSD Spatio-Spectral Decomposition

SSEPs Steady Sstate Evoked Potentials

SSVEPs Steady State Visual Evoked Potentials

SVM Support Vector Machine

Chapter 1

Introduction

A Brain-Computer Interface (BCI) system is defined as a communication system conveying messages through brain activity without any muscle movement [1–4]. The first BCI system which was invented in 1973 [5], used Electroencephalography (EEG) for communication purposes. BCIs have been successfully used in clinical applications over the past years. As an example, locked-in patients communicated with their environment via a BCI speller device [6–8]. Moreover, BCI was used as an assistive device for partially disabled people [9–11]. BCI technologies have also proven effective in helping the physically disabled control a wheelchair using only their brain signals [12–14]. In recent years, BCI has been also used for stroke rehabilitation [15–17]. Although the ultimate goal of BCI is to provide a communication tool for paralyzed or disabled people [18, 19], BCI has other non-medical applications as well [20–23].

A BCI system records the brain activity either invasively or non-invasively [7]. However, invasive-BCIs are less popular compared to non-invasive BCIs. Among several non-invasive techniques for recording brain activities, EEG is mostly used due to its good temporal resolution, low cost, and portability. An EEG-based BCI system can be controlled by the modulation of EEG rhythms. Motor Imagery (MI) or imagination of movement is one of the mental tasks that induces changes in EEG rhythms and thus can be used for controlling an EEG-based BCI [24]. Numerous signal processing

and machine learning algorithms have been developed to detect MI from the recorded EEG [25] and translate it into a command for controlling an application such as moving a cursor on a computer screen [26–28].

1.1 Research motivations

Over the years, BCI research breakthroughs and technology developments have significantly improved the lives of patients with severe motor disabilities and impairments [28–30]. Current BCI systems are scarcely used in out-of-lab scenarios and hence they need to be improved in various areas such as signal acquisition techniques, signal processing methods, machine learning algorithms, hardware development, and usability [19–21, 31, 32].

The main motivation of this thesis is to predict and improve the performance of EEG-based MI-BCI systems in order to make it a more practical technology. Ideally any user should be able to control a BCI system. However, several years of MI-BCI research have demonstrated that the performance of subjects vary considerably [33, 34]. Some subjects have poor performance and thus cannot use a BCI system [34]. The reasons for the performance variation of MI-BCI users have not been studied extensively.

One of the current issues in EEG-based MI-BCI systems which can adversely affect the performance of the subjects is BCI deficiency, also known as BCI illiteracy. BCI deficiency is defined as the inability of subjects to use a BCI system [34, 35]. It has shown that some subjects cannot modulate their brain signals [33]. BCI performance predictors [35–38] can help detect the poor performance subjects. However, none of the proposed predictors have been used in practical BCI applications to improve performance of the subjects.

Poor performance can also be a subject of the machine learning algo-

rithms employed in EEG-based BCIs. Normal machine learning algorithms cannot intuitively handle the non-stationarity of the EEG, which is still one of the main issues in EEG-based BCI systems [34, 39, 40]. Due to non-stationarity, the statistical characteristics of EEG signal may vary over time, and thus the feature space from calibration session into evaluation session may be changed. In other words, the classifier trained based on the calibration session is no longer optimal for the evaluation session. Adaptive machine learning algorithms such as adaptive classification methods [41, 42] or adaptive feature extraction methods [43, 44] can address the non-stationarity of EEG data, and accordingly enhance the performance of the MI-BCI users.

1.2 Research objectives

This study mainly focuses on addressing some of the current issues in EEG-based MI-BCIs to improve the performance of the users. BCI performance reflects how well a BCI system is controlled by a user, therefore improving the BCI performance will yield to a more practical BCI system. To do so, we can help subjects with BCI deficiency improve their performance, and develop adaptive algorithms. BCI deficiency is one of the reasons of having poor performance. The neurophysiological reasons behind BCI deficiency and performance variation across subjects need to be investigated. The performance predictors previously proposed to detect BCI deficient subjects have not been used in practical applications. To the best of our knowledge, no BCI experiment has been designed to specifically help poor performance subjects enhance their BCI performance. Non-stationarity of EEG signal, on the other hand leads to BCI performance mitigation. Therefore, developing adaptive feature extraction or classification algorithms which reduce the difference between sessions can improve the BCI performance.

Accordingly, the two main objectives of this thesis are defined as follows:

- ***Detecting subjects with MI-BCI deficiency and proposing a method to improve their BCI performance:***

  The first objective of this thesis is to propose novel performance predictors to estimate the performance of the subjects. In other words, we aim to detect subjects with BCI deficiency, and investigate the neurophysiological reasons behind MI-BCI performance variations. By doing so, we can design a novel experiment to help subjects improve their performance and thus better control a BCI system.

- ***Improving EEG-based MI-BCI performance through addressing inter-session non-stationarity:***

  The second objective of this thesis is to propose novel adaptive algorithms to address inter-session non-stationarity and thereby improve the MI-BCI performance. Adaptive classification and feature extraction methods can address inter-session non-stationarity.

The analysis in this thesis are applied to three datasets recorded in the Neural Signal Processing laboratory of Institute for Infocomm Research (I^2R), Agency for Science, Technology and Research (A*STAR), Singapore. The experiments at I^2R were carried out in accordance with the criteria approved by the Institutional Review Board of the National University of Singapore (NUS). Another big dataset used in this thesis was jointly recorded by the Neurotechnology group at Berlin Institute of Technology (TÜ Berlin) and Tübingen. The recorded data in Berlin and Tübingen were approved by the Ethical Review Boards of the Medical Faculty, University of Tübingen.

1.3 Outline of the thesis

This thesis is organized into eight chapters. The background and basic ideas about the design of a BCI system are explained in Chapter 2. Chapter 3 and 4 propose two novel performance predictors to estimate the performance of MI-BCI users. Chapter 5 describes our proposed experiment to improve the performance of MI-BCI users. Our proposed algorithms to address inter-sessions non-stationarities are described in Chapter 6 and 7. Some parts of this thesis are based on our own previous publications in [45] (Chapter 3), and [46] (Chapter 4), [47] (Chapter 5), [48, 49] (Chapter 6). The detailed contents of each chapter are listed as follows:

- **Chapter 2** gives an overview of BCI and explains some of its applications. This chapter also reviews some of the commonly used methods for recording brain activity, signal processing and machine learning methods for analyzing brain signals. Moreover, some of the BCI applications and their current issues are also discussed in this chapter.

- **Chapter 3** proposes a novel neurophysiological coefficient to predict the performance of MI-BCI. The new proposed coefficient is computed from the spectral power of EEG rhythms over different regions of the brain. The predictor can successfully detect low performance subjects.

- **Chapter 4** presents another neurophysiological performance predictor based on the spatio-spectral information in the EEG signal. This new proposed predictor is computed based on spatial and spectral characteristics of the resting state EEG and pre-cue of feedback runs. The predictor can successfully estimate the performance of a large group of novïce subjects.

- **Chapter 5** proposes a new experimental design to improve the MI-

BCI performance of the subjects through neurofeedback training of resting state alpha using the proposed predictor in Chapter 4.

- **Chapter 6** includes a novel method based on Kullback-Leibler (KL)-distance weighting to adapt features extracted using Filter Bank Common Spatial Pattern (FBCSP), to address inter-session non-stationarity of EEG data and improve the MI-BCI performance of the subjects. This chapter explains the benefits of calibrating MI-BCI with passive movement.

- **Chapter 7** describes a novel adaptive classification algorithm, Adaptive Extreme Learning Machine (A-ELM), to address inter session non-stationarity. The proposed algorithm used limited number of data from the evaluation session to adapt the initial classifier.

- **Chapter 8** concludes this thesis and provides a summary of our proposed methods.

Chapter 2

Brain-Computer Interface

2.1 Introduction

According to the definition given by Wolpaw et al. in the first international BCI meeting [50], BCI is a system provides a new communication and control channel for its users, which is not the normal output channel of peripheral nerves and muscles. A BCI system extracts the brain patterns associated with recorded brain activities and thus helps users communicate with the external environment.

Many aspects of BCI systems are currently being investigated in order to improve the practical prospects of design in BCI systems. Research areas include evaluation of invasive and noninvasive technologies to measure brain activities, evaluation of control signals (i.e. brain activity patterns used for communication), development of algorithms for translation of brain activity patterns into computer commands, and development of new BCI applications. This chapter briefly reviews the main aspects of a BCI system and highlights some of the recent developments and current issues. More complete reviews can be found in [1, 2, 50, 51].

Typically, a BCI system contains three main components which are shown in Figure 2.1. The first component is a signal acquisition unit which records the brain activity of the user through appropriate sensors. Some of the commonly used signal acquisition techniques are reviewed in Section 2.2. An ideal signal acquisition technique requires to have minimum side

effects on the user, be cheap and also easy to use. More importantly, the signal acquisition technique should have high temporal and spatial signal resolution.

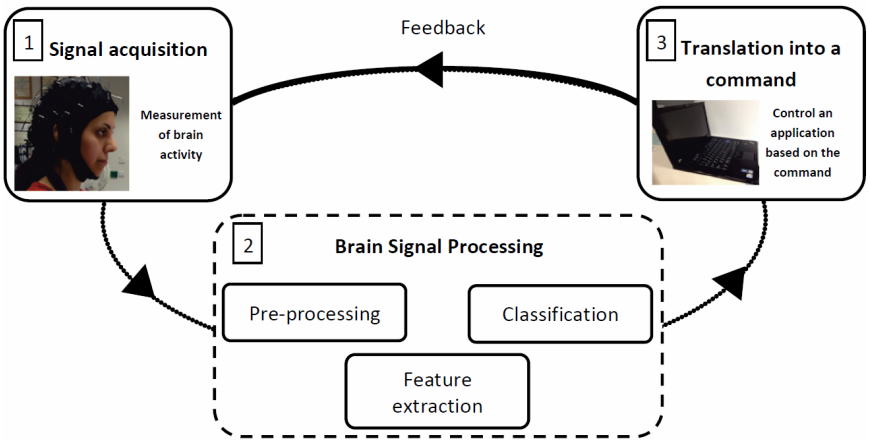


Figure 2.1: General architecture of an online BCI system. BCI has typically three main components which are responsible for: 1) acquiring brain activities, 2) processing the recorded signal, 3) translating the identified mental state into command for controlling an application. A feedback is also provided in online scenario. A calibration phase is needed before online scenario which is not shown in this figure.

The second component of a BCI system is a brain signal processing unit. The recorded brain signals need to be preprocessed to remove the noise and artifacts. At the next step, the features are extracted from the denoised signal and fed into a classifier to identify the mental state of the user. More information about the brain signal processing methods are provided in Section 2.3.3. Since the focus of this thesis is on EEG-based BCI, only those methods which are commonly used in EEG-based BCI are reviewed.

The third component of a BCI system is responsible for translation of the identified mental state or the output of the classifier into a command. The derived command is used for controlling an application. Finally, based on the identified mental state a feedback is provided to guide the user better modulating his brain signals. However, there are several debates about the role of feedback. More details are provided in Section 2.3.4.

The ultimate goal of BCI system is to be used in real world applications such as therapeutic applications. Some of the most practical applications of EEG-based BCI are reviewed in Section 2.4. Despite many years of research, there are still some issues limited possible applications of BCI systems. Some of the current BCI issues which are still under investigation are discussed in Section 2.5.

The BCI system introduced in Figure 2.1 needs to be calibrated before being used in an online scenario. Therefore, generally in most BCI experiments, a short calibration session is recorded to train the model. The trained model is used later in an evaluation session. However, in order to have a more practical BCI system, the calibration time needs to be minimized [52, 53]. Subject independent classifiers have been also proposed to reduce the calibration time of the BCI experiment [54, 55]. Such classifiers are trained based on the data collected from other subjects. However, they may not be as accurate as subject specific classifiers which are typically used in BCI experiments.

2.2 Signal acquisition techniques

The first component in all BCI systems is signal acquisition unit which is a neural interface component. This component is a hardware device which detects and records the brain activity. Several methods have been used for acquiring brain signals. The signal acquisition technique is chosen based on the intended use of a BCI system and its target users. Generally, the chosen technique must be safe for the user, and its invasiveness must not be more than necessary. The technique should be also capable of recording sufficient information from the brain to control the BCI. In light of these requirements, several technologies have been developed. The signal acquisition techniques

in BCI systems can be categorized into invasive and non-invasive methods. Each of these techniques is more suitable for a specific study and application.

2.2.1 Invasive signal acquisition

The invasive methods have some clinical risks and thus have been rarely used in real applications [7, 56, 57]. Electrocorticography (ECoG) is one of the invasive modalities which records field potentials from set of neurons [58]. The ECoG electrodes are placed directly on the surface of the brain through a surgery. Another invasive technique involves inserting microelectrode arrays into the brain for acquiring electrical brain activities [58]. Similar to ECoG this approach also needs surgery to insert electrodes into the brain. The microelectrode arrays record neural action potentials (spikes) from each neuron and/or local field potentials from set of neurons. However, this method is highly invasive to the brain tissue and the recorded signal is unstable and deteriorated over time, since the brain tissue reacting electrodes as foreign objects.

2.2.2 Non-invasive signal acquisition

Contrary to invasive signal acquisition techniques, non-invasive methods are widely used in BCI applications [59]. Clinical risks are minimized in non-invasive methods. Different non-invasive modalities have been proposed over the past years. The intended application of the BCI system determines which modality is more suitable for acquiring the brain activity. The acquired signal by different modalities has different spatial and temporal resolution.

Magnetoencephalography (MEG) is one of the non-invasive methods which records the magnetic fields produced by electrical currents of the brain. Only a few research groups are working on BCI controlled by MEG. Since the device is too expensive, bulky and cumbersome, it is not considered the most

suitable technique in many applications or in out-of-lab scenarios. Functional Magnetic Resonance Imaging (fMRI) is another non-invasive technique which captures Blood Oxygen Level Dependent (BOLD) of brain. The blood flow of the brain is coupled with neural activation inside the brain. However, fMRI similar to MEG is too bulky and not appropriate for out-of-lab scenarios. Functional Near-Infrared Spectroscopy (fMRI) is a functional neuroimaging technique which measures the hemodynamic or BOLD responses associated with neural activity of the brain. Contrary to MEG and fMRI, fNIRS can be used in out-of-lab scenarios. However, the temporal and spatial resolution of fNIRS compare to other modalities is not high enough. Another non-invasive method for recording brain activity is EEG. EEG records the electrical activity of brain from scalp. Comparing to other modalities, EEG is the most commonly used signal acquisition technique in BCI through many years. Figure 2.2 compares the spatial and temporal resolution of non-invasive signal acquisition techniques.

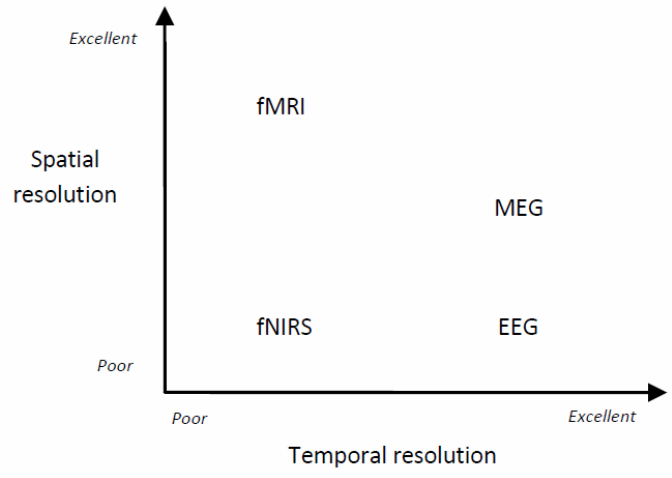


Figure 2.2: Comparing spatial and temporal resolution of the signal acquisition techniques used in non-invasive BCI. The temporal resolution varies from 0.05 to 1 second, and the spatial resolution is between 1 millimeter and 10 millimeter.

As shown in Figure 2.2, MEG has high temporal and moderate spatial

resolution, fMRI has high spatial resolution but low temporal resolution, and fNIRS comparing to the other methods has lower spatial and temporal resolution. EEG has high temporal resolution which is a necessity for real-time BCI. Although spatial resolution of EEG is relatively low, its low cost and portability makes it practical for many applications.

2.3 EEG-based Brain-Computer Interface

EEG is the mostly used modality for recording brain activity in BCI systems. Over the years, different parts of EEG-based BCI systems have been developed to make them more suitable for real-world applications. Designing proper sensors for recording EEG is still one of the current issues in BCI systems.

EEG records the brain electrical activity from electrodes placed on the scalp [60]. Typically, Ag/AgCl electrodes are used for recording EEG which need to be used with conductive-gel. Another type of electrodes has been emerged named "dry electrode" [61]. The main advantage of the dry electrode is its convenience to use. Dry electrodes are mostly suitable for long-term EEG recordings [61]. Despite the improvements over the past few years, the quality of the recorded signals from dry-electrode technology is still affected by sweat on skin, movement, and environmental noise. Therefore, they are rarely used in practical applications of EEG-based BCIs.

In EEG recording, an EEG cap is used on which the Ag/AgCl electrode locations are fixed, so that the locations are not changed during each session of BCI experiment. The EEG electrode names and locations are usually assigned according to the international 10-20 system. This system assigns standardized names to the electrodes. The locations of the electrodes on the scalp are also determined in 10-20 system. The electrode names stand for

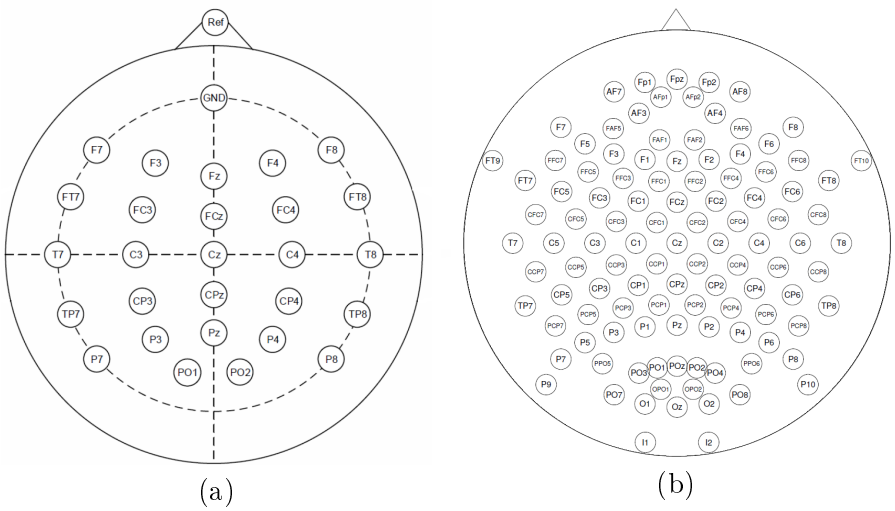


Figure 2.3: The EEG electrode locations on the scalp based on the international 10-20 system. Two different montage with (a) 27 channels, (b) 118 channels are shown. Left hemisphere electrodes have odd numbers, and the electrodes on the right hemisphere have even numbers, and the numbers increase from the center to the outer electrodes. The electrode names represent the underlying cortical regions: F is frontal, P is parietal, O is occipital, T is temporal, C is central.

the underlying cortical regions. The electrodes located in the left hemisphere have odd numbers, and even numbers are assigned to those electrodes located in the right hemisphere. In this thesis, two different montages had been used: 27 channels and 118 channels. Figure 2.3 shows the electrode montage using 27 channels (a) and 118 channels (b).

2.3.1 EEG rhythms

The EEG signal recorded from the scalp is a superposition of neural oscillations. Studying neural oscillations' origins and roles is one of the key research topics in neuroscience [62–64]. The neural oscillations or rhythms are denoted by Greek letters. Table 2.1 briefly introduces these rhythms. The EEG rhythms in a certain frequency band have a special location and function. The spatial location and neurophysiological role of the rhythms are also reported in the Table 2.1. The α, μ, and β rhythms have prominent role in EEG-based BCIs.

Table 2.1: Overview of the EEG rhythms

Rhythm	Frequency (Hz)	Location	Characteristics
δ (Delta)	0.1-4	Frontal and posterior regions	Sleep study
θ (Theta)	4-7	various locations	Attention analysis, sleep study
α (Alpha)	8-13	Parietal, occipital and temporal regions	Motor activity, Visual attention, Resting state analysis, Wakefulness and drowsiness analysis
μ (Mu)	8-13	Sensorimotor cortex	Motor activity
β (Beta)	13-30	Frontal and somatosensory cortex	Motor activity, Attention and Engagement study
γ (Gamma)	>30	Various locations	Motor activity, Emotion recognition, Visual and spatial attention

Alpha rhythms

Alpha waves are the neural oscillations in the frequency range 8-13 Hz. They usually appear during relaxation with eyes open, wakefulness, drowsiness and sleep [60, 65]. The Alpha rhythms can be measured in parietal, occipital and posterior temporal regions of the brain. In fact, alpha rhythms can be originated in different cortical areas [66]. Depending on the desired functional task, the recorded alpha waves measured from various locations of the brain may be studied.

Typically, the amplitude of alpha rhythms in eyes open condition are lower than the amplitude of alpha rhythms in eyes-closed condition [67]. Alpha waves suppress in task relevant regions and enhance in non-relevant brain regions [68]. Alpha rhythms have also been proven to correlate with several mental tasks including attention, consciousness, and perception. In general, alpha suppression requires external attention such as visual stimuli [69]. In contrast, alpha is enhanced during mental tasks such as short- and long-term memory task and working memory task [70]. The suppression and

enhancement of alpha rhythms, and their topographic distribution mostly depends on the task performed by the subject. In addition, the frequency range of alpha waves are also subject specific [71].

Mu rhythms

Mu or "Rolandic alpha rhythms" are EEG oscillations appearing in the frequency range of 8-13 Hz over sensorimotor cortex. There is also another peak frequency in the beta band around 20 Hz named "Rolandic beta rhythm" [72, 73]. The location of these rolandic rhythms and their modulation by actual limb movement and somatosensory input proved a functional relationship between these rhythms and sensorimotor system. Moreover, the motor cortex excitability depends on the strength of these rhythms [73].

Although the frequency range of alpha and mu rhythms are almost similar, they have different functionality and topography. Mu rhythms usually do not show symmetry between the two hemispheres. Mu rhythms had shown to be attenuated during motor execution [74], motor imagination [75], and motor observation [76, 77], while less affected by visual stimulation [73].

Beta rhythms

Beta waves are oscillation in the range of 13-30 Hz associated to attention, engagement, concentration and also motor activity tasks. Rolandic beta rhythms are associated with the motor cortex, while mu rhythms are related to the somatosensory areas. The amplitude of beta oscillations are generally lower than mu oscillations. In fact, the amplitude of EEG rhythms are negatively correlated with the frequency of the rhythms [78]. Similar to mu rhythms, beta rhythms over sensorimotor area are also decreased during the motor execution, motor imagination, and motor observation. However, beta rhythms have shown to be enhanced after the motor task (please refer to

Section 2.3.2.3).

2.3.2 Control signals used in EEG-based BCI

Neurophysiological signals or brain activity patterns used to drive EEG-based BCI systems can be categorized into two types: 1) Evoked Potentials (EPs), and 2) spontaneous signals. In the following these two types of signals are briefly reviewed.

2.3.2.1 Evoked potentials

The EPs are automatically generated by the brain in response to a stimulus, thereby users do not need any specific training. However, these control signals are uncomfortable and tiring for the users since they require external stimulus. Two of the main examples of EPs include Steady Sstate Evoked Potentials (SSEPs) [79] and P300 [6, 13, 80].

Steady state evoked potentials

SSEPs are brain activities generated in response to periodic stimuli. Analyzing the EEG power spectral demonstrates a rhythmic increase in the power of the EEG signals at stimulation frequency, its harmonics or sub-harmonics. Steady State Visual Evoked Potentials (SSVEPs) have been used more in comparison to the other types of SSEP control signals [79]. SSVEP appears mostly in visual areas, since a visual stimulus is provided for the user. Other SSEPs such as somatosensory SSEP [81] or auditory SSEP [82] have been also used for controlling an EEG-based BCI system.

P300

The P300 wave is elicited from EEG approximately 300 ms after unexpected stimuli. This potential is mainly generated when using "odd-ball" paradigm.

The users should pay attention to frequent and non-frequent target items. Typically, the appearance of non-frequent items triggers a P300 which is usually located in parietal area. The EEG-based BCI systems controlled by P300 have been extensively used in helping patients who suffered from Amyotrophic Lateral Sclerosis (ALS) [80]. The most well known application of P300 for ALS patients is speller device [8]. P300 has been also used for internet browsing [83], controlling a wheelchair [13] and also as an assistive devices [84].

2.3.2.2 Spontaneous signals

The other group of control signals used to drive an EEG-based BCI system are spontaneous signals. Table 2.2 compares some of the important advantages and disadvantages of spontaneous signals and EPs. Spontaneous signals are intuitive and natural to use without any external stimulus. Therefore, they might be less tiring for the users in comparison with EPs. However, the BCI users need some training to control a BCI system with spontaneous signals. Advanced machine learning and signal processing methods reduce the need of subject training [85].

Table 2.2: Comparing the neuro-physiological signals used to drive an EEG-based BCI system. Two examples of control signals along with some of their advantages and disadvantages are listed in the table.

Control Signal	Examples	Advantages	Disadvantages
Evoked potentials	- SSEP - P300	- High information transfer rate - No learning	- No user's attention - Require external stimulus
Spontaneous signals	- SMRs - SCPs	- Intuitive and natural to use - No external stimulus	- Require learning - Lower information transfer rate than evoked potentials

Sensorimotor Rhythms (SMRs) [86–88] are undoubtedly the most commonly exploited signals in the category of spontaneous signals. Due to their

importance, they are introduced in more details in Section 2.3.2.3. Slow Cortical Potentials (SCPs) [89] also fall within the category of spontaneous signals to drive EEG-based BCIs.

Slow cortical potentials

SCPs are the slow cortical activity changes in the frequency range of 1-2 Hz. SCPs are stable signals that last from hundreds milliseconds to several seconds [90–92]. Users need to generate negative and positive SCPs to control a BCI system. As an example, increasing or decreasing excitation is one of the possible methods for generating SCPs. However, controlling SCPs need a long term training which is one of the main limitations of the BCI systems controlled by SCPs.

2.3.2.3 Sensorimotor rhythms

Sensorimotor rhythms (SMRs) are in the category of spontaneous signals and mostly used to drive EEG-based BCI systems. SMRs are defined as brain rhythms in mu (8-13 Hz) and beta (13-30 Hz) frequency bands over sensorimotor areas. During motor behaviours the rolandic rhythms or SMRs decrease, which is known as event-related desynchronization (ERD) [78, 93, 94]. This decrease is due to an internally or externally paced event such as a voluntary movement. An increase of SMRs known as event-related synchronization (ERS) may occur after the motor task, which is typically observed in beta band [78]. Two strategies have been proposed to control these sensorimotor rhythms:

Operant conditioning

Operant conditioning [26, 95] is one of the earliest methods in BCI applications. In operant conditioning-based BCI, the user performs a mental

strategy which is comfortable for him. A fixed translation algorithm is used to generate a feedback signal from EEG. Feedback guides the user to better modulate his brain activities and control the BCI system. However, this method needs a very long training time from several weeks to even several months [26] which is not very desirable specially for the users.

Motor Imagery (MI)

MI can be defined as a dynamic state during which the representation of a specific motor action is internally reactivated within working memory without any actual motor movement [96]. MI induces changes in SMRs over sensorimotor areas [24]. Several imaginary actions can be used for modulating the SMRs. It has shown that MI of hand [16], foot [97], and swallow [98] can control an EEG-based BCI. MI in comparison with operant conditioning needs less training time. Therefore, it is more desirable in real applications. In this thesis, we have focused on EEG-based BCI systems which controlled by MI.

MI can be performed in different ways, some of the common paradigms are: visual versus kinesthetic, first or third person imagination [99]. Visual MI is the mental image of the movement or in simple words a subject imagines how is the movement looks like, while kinesthetic MI is the proprioceptive aspects of the movement. Typically, kinesthetic imagery performed in first person perspective which means the subject imagines his own movement. Third person MI is not easy for some subjects since they have to imagine watching movement of another person. All data used in this thesis are collected from subjects while performing first person kinesthetic MI.

2.3.3 EEG processing

The recorded EEG signals need to be processed and translated into commands to be used for controlling a computer-based application. EEG processing includes preprocessing, feature extraction and classification which are briefly explained in the following sections.

2.3.3.1 Preprocessing

The acquired EEG signal is typically noisy, and contaminated by artifacts such as muscular artifacts or eye movements. Preprocessing enhances the signal to noise ratio of the recorded EEG signal by removing the irrelevant information from EEG. Some of the preprocessing methods are explained as follows:

- **Temporal filters**

  The neurophysiological signals are extracted from a specific frequency band. Temporal filters such as band-pass filters are generally used to restrict our analysis to the frequency band of interest. As an example, generally in MI-BCI the ERD/ERS can be found over mu (8-13 Hz) and beta (13-30 Hz) bands [27, 78], thereby the acquired EEG is band-pass filtered in 8-30 Hz frequency band.

- **Artifact removal**

  Artifacts are undesirable and can easily affect the extraction of neurophysiological signals. They may be generated due to different reasons such as: movement of body parts like, eyes, tongue, arms or change in skin resistance (i.e., sweating). Independent Component Analysis (ICA) is one of the effective tools for artifact removal in BCI [100, 101]. In the most BCI experiments, subjects are typically instructed to minimize their body limb movement. Moreover, the collected EEG data

are visually inspected and those EEG trials with excessive artifacts are removed. The other sources of interferences, such as power line interference, can be usually removed by the EEG hardware device.

- **Spatial filters**

  Similar to temporal filtering, spatial filtering aims to reduce the effect of irrelevant information embedded in the EEG. A spatial filter gives weights to electrodes, less weights are assigned to the electrodes which are irrelevant to the targeted task. Common Average Reference (CAR) [102] and the Surface Laplacian (SL) [33] are two of the simplest spatial filters used in EEG processing. CAR removes the average activity of the other electrodes from each electrode, while SL removes the average activity of the neighboring electrodes. The most commonly used spatial filter is Common Spatial Pattern (CSP) [88, 103, 104]. CSP spatially filters EEG signal to maximize the discrimination between the classes involved.

- **Other preprocessing algorithms**

  Some other such as Principal Component Analysis (PCA) [105] or Domain Space Adaptation (DSA) [106] has been also used in BCI applications to enhance extraction of neurophysiological signals.

2.3.3.2 Feature extraction

Numerous feature extraction techniques have been studied and proposed for BCI applications [107]. Three main categories of features are described below:

- **Temporal features**

  Temporal features are those defined based on the temporal variations

of the signals. The amplitude of raw EEG signals [80], auto-regressive parameters [108, 109] are the commonly used temporal features.

- **Spectral features**

  Performing a mental task induces changes in neurophysiological signals. As an example, MI results in ERD/ERS. These changes can be captured from the power of EEG rhythms over specific frequency bands. The spectral or band-power features are commonly used in BCI applications [15, 88].

- **Temporal-spectral features**

  Due to significance of both temporal and signal frequency information in BCI, temporal-spectral features are introduced to consider both sources of information into account. Short-time fourier transform [110] or wavelets [111] are the two examples of temporal-spectral features.

2.3.3.3 Classification

The extracted features from EEG signal are fed into a classifier. So far, several different classifiers have been applied in BCI applications [112]. However, linear classifiers such as Support Vector Machine (SVM) [43, 47, 113–116] and Linear Discriminant Analysis (LDA) [41, 42, 46] are the most popular linear classifiers. Some nonlinear classifiers such as Naïve Bayesian Parzen Window (NBPW) has been also used in BCI applications [117, 118].

2.3.4 Feedback

A feedback in a BCI system informs a subject about his performance. The role of feedback may vary across subjects. McFarland et al. in [119] showed that feedback can have both positive and negative effect on modulating EEG rhythms. As an example, during MI a subject may receive wrong feedback,

while correctly imagining the movement. This may lead the user to change his strategy of imagination and finally decrease the performance. Therefore, providing feedback is not always suggested [120].

2.4 EEG-based MI-BCI applications

Over the past few decades, motor imagery (MI) has been widely used for controlling EEG-based BCI in various applications. In this section some of the applications of the EEG-based MI-BCI, which is the focus of this thesis, are reviewed.

Neurorehabilitation

Motor impairment after stroke is the most important reason of permanent disability [96]. During the last decade, several methods were developed to support stroke rehabilitation [96, 121–124]: Active Movement Training (AMT), electromyography biofeedback, robotics, and mental practice with MI. Currently, there are enough evidences showed that imagination of movement has positive role on rehabilitation after stroke [96]. MI in combination with physical therapy leads to enhanced motor outcomes for stroke survivors.

Motor system has shown to be activated similarly during MI and actual movement [9]; therefore, MI would be useful approach for subjects with different levels of disabilities. However, there are unsolved issues associated with MI: it is hard for a therapist to evaluate the performance of the MI performed by a subject and also a subject himself has no feedback about his own performance. BCI system can help to overcome these problems [10, 122]. In fact, MI-BCI is known as a possible approach to functional recovery after stroke. MI-BCI has been used for both upper limb [15–17, 96] and lower limb therapy [97, 125, 126].

Other medical applications

Some of the patients with severe disabilities such as ALS or completely locked-in patients lack communication ability with their environment. Controlling a cursor on the monitor using MI-BCI was a promising method which provided the patients a new communication pathway [26, 127]. MI can be a control signal for controlling a wheelchair. It has shown that a subject with a Tetraplegic can control a wheelchair in virtual environment [128]. MI can be used for controlling an orthosis to open and close a paralyzed hand of a subject [129].

Non-medical applications

Although cursor control and speller were initially introduced to help patients with motor disabilities, they can be also used by healthy subjects in some entertaining applications. Using BCI for gaming or virtual reality purposes is quite interesting for healthy subjects. As an example, virtual and real helicopter control are the two interesting applications of the MI-BCI [130–132].

2.5 Current issues in EEG-based MI-BCI

Despite of development in signal processing and machine learning algorithms over the past few years, BCI systems still face some issues such as:

***BCI deficiency*:**

One of the fundamental issues in BCI is that around 15% to 30% of subjects are not capable of using a BCI system [133], and their BCI performance is relatively low. This phenomenon is known as BCI deficiency or illiteracy [33]. One of the possible reasons of BCI deficiency is related to the functionality of

the brain. It seems that subjects BCI deficiency are unable to modulate their brain rhythms [45]. However, the possible reasons of such inability have not been extensively studied. The other reason lies within the machine learning used for detecting the brain patterns [134]. It is quite possible that the brain waves have been modulated, but this modulation cannot be detected by machine learning algorithm. Non-stationarity of EEG signal caused a difference between feature space of the two sessions [47]. Therefore, the employed machine learning algorithms are no longer appropriate to handle the transition from calibration to evaluation session.

***Non-stationarity of EEG*:**

The EEG signal in response to a stimulus or during a specific task differs from one subject to another. Moreover, the EEG signal differs from trial to trial and from day to day. This change in EEG signal properties over time is known as non-stationary of EEG which is one of the major challenges in EEG-based BCIs [39]. Non-stationarity of EEG occurs due to different factors such as changing the subject states or moods, artifacts, or even changing the EEG cap from one session to another session [48].

The non-stationarity of EEG signal can be seen in both inter- and intra-session [40, 106]. Analyzing the feature space of the calibration and evaluation session demonstrates that there might be a shift inside the feature space, therefore the trained model based on the calibration data might not be optimal for the evaluation session anymore.

***Other issues*:**

Apart from BCI deficiency and EEG non-stationarity, there are some other challenges in MI-BCI which are not the focuses of this thesis. EEG signals are usually noisy and can be affected by artifacts such as eye blinking or muscle

movements, therefore suitable preprocessing methods should be applied to remove noise and artifacts. A calibration session which is typically required in MI-BCI systems is time-consuming for both user and operator and it is desirable to minimize or eliminate the BCI calibration time [52, 85]. Finally, EEG-based MI-BCI systems are scarcely used in out-of-lab scenarios. One reason could be could be due to use of conductive gel on the scalp during setup, which causes inconvenience to the user.

2.6 Summary

Imagination of movement or MI is one of the brain activities which can be used to control an EEG-based BCI system. EEG-based MI-BCI has shown effective in many medical applications; even in subjects with sever motor impairment.

Despite the improvement in BCI technology, signal processing and machine learning algorithms, BCI systems are not yet fully equipped to be used in out-of-lab scenarios. There are still several issues regarding EEG-based BCI systems. One of the important challenges is BCI deficiency. It had shown that not all the subject can use a BCI system, understanding the reason of such inability may help us to have a novel design of a BCI system which can help these poor performance users control a BCI system.

Another important issue is non-stationarity of EEG signals. The statistical characteristics of EEG signal varies over the time, which makes the initial calibrated model to be suboptimal for latter sessions. Addressing this issue can lead to improve the BCI performance.

Chapter 3

Predicting the performance of MI-BCI using a novel neurophysiological coefficient

---

3.1 Introduction

EEG-based MI-BCI systems have been widely used for both therapeutic and non-therapeutic applications [15, 21]. However, one of the current issues in most MI-BCI systems is that a non-negligible number of users (15 to 30%) cannot perform MI well [134]; hence they cannot properly use BCI systems. In fact, there is a big variance in the performance of BCI users [35]; however, the reason of why some users are BCI deficient is still under investigation by many researchers.

Performing MI results in ERD/ERS of EEG rhythms [78]. In other words, stronger ERD/ERS can indicate how well the subject performs MI task. BCI deficiency in subjects using MI-BCI can be possibly attributed to their inability in modulating EEG rhythms [134]. It may be also because of a mismatch between calibration session and evaluation session due to non-stationarity in the EEG [40]. Machine learning algorithms used in BCI systems may not be capable to deal with such non-stationarity, which may adversely affect the BCI performance in the evaluation session [134]. Moreover, the reasons of

BCI deficiency may vary among subjects.

The first step towards a cure for BCI deficiency is to detect the potential BCI deficient subjects. Performing a BCI experiment is quite time consuming, therefore it is highly preferable to identify the poor performance subjects, prior to the whole BCI experiment. Having some prior knowledge about the performance of the subjects may lead us to investigate other possible reasons of performance variation in different subjects and also yield in designing a novel experiment which aimed to help users with BCI deficiency [35, 38].

Defining a performance predictor to predict performance of BCI users is highly valuable. The predictor may estimate the performance of BCI users without performing a long time experiment. It can be effectively used in clinical applications, as an example it can quickly help us to investigate whether the BCI system is an appropriate assistant device for a patient. Although several predictors have been proposed over the past few years, so far none of them has been widely used in BCI experiments. To the best of our knowledge, no study has effectively used the BCI performance predictors to help the subjects better control a BCI system.

In this chapter, a novel neurophysiological coefficient is proposed to predict the classification performance of MI-BCI. The proposed coefficient is computed from the spectral power of pre-cue EEG rhythms over different regions of the brain. The feasibility of predicting the classification performance of the MI-BCI users from the proposed coefficient is further investigated. We assume that there is a correlation between our proposed EEG rhythm-based coefficient and the performance of the users. In comparison with previously proposed predictors, this coefficient captures spatial information from the brain in addition to spectral information.

The remainder of this chapter is organised as follows: Some of the previ-

ous performance predictors are reviewed in Section 3.2. Section 3.3 describes our proposed predictor, followed by experimental setup in Section 3.4. The performance evaluation method used in this chapter is explained in Section 3.5. Section 3.6 and 3.7 present the experimental results and discussion. Finally, the summary of this chapter is brought in Section 3.8.

3.2 Neurophysiological versus psychological performance predictors

The previous studies on BCI performance prediction can be categorized into two different groups: the first group focused on modulation of slow cortical potentials [91, 92, 135, 136], while the second group proposed psychological/ neurophysiological predictors for SMR-based BCIs which are the main focus of this thesis [35, 37, 38, 86, 137–139].

Psychological parameters have shown to have moderate but meaningful role on BCI performance [37, 86, 138]. Several psychological parameters such as visuo-motor coordination, attention, personality or motivation were studied in [138]. However, it has been shown that only the visuo-motor coordination and concentration ability of the users can be considered as psychological predictors for SMR-based BCIs. They can be tested by Two-Hand Coordination Test (2HAND) and Attitude Towards Work (AHA) test, respectively. It has shown in [86] that mood and motivation affect the ability of the subjects in learning how to use a BCI system. Hence, these psychological predictors, which reflects the users' feelings, can predict the BCI performance. However, some of the psychological parameters (e.g. personality test B5PO [138]) are based on self-assessment criteria and not well quantified. Therefore they may not be reliable for predicting the performance of the users.

The second category of the performance predictors proposed for SMR-

based BCI are neurophysiological predictors which are defined based on the EEG rhythms. A neurophysiological predictor was proposed in [35] based on the μ- (9-14 Hz) and β- (20-30 Hz) rhythms over sensorimotor area. They found positive correlation, r=0.53 between their proposed predictor and the BCI performance of the subjects. In a more recent paper, the alpha band activity extracted from 1 second pre-stimulus EEG data has been introduced for predicting the performance of EEG trials. It has shown, higher alpha band activity prior the cue resulted in significantly better classification performance [139]. Gamma band power was also proposed as another performance predictor for SMR-based BCI [38, 137]. It has shown that γ oscillation has causal influence on SMR, thus can be used for predicting the performance of SMR-BCI.

The pre-stimulus alpha [140–145] and theta [146, 147] have been used for predicting the performance of some other mental tasks rather than MI. However, the reported effects vary according to the performed mental task. In other words, for some mental tasks such as visual perception lower alpha results in higher performance, while for some other tasks such as cognitive or memory tasks higher alpha leads to better performance.

Considering all previously mentioned studies, it can be concluded that pre-stimulus EEG contains useful information about the task outcome performance. In other words, the state of the brain before providing a stimulus has a role on the performance of the subject. Therefore, it can be assumed that by knowing the state of brain over pre-stimulus time segment, the task performance of the user can be predicted.

3.3 Proposed neurophysiological coefficient for performance prediction

In this chapter, we aim to propose a novel neurophysiological coefficient to predict the classification performance of MI-BCI from pre-cue EEG data. The state of subject's brain preceding a cue or stimulus can be defined using EEG rhythms. In some Attention Defcit Hyperactivity Disorder (ADHD) studies, the ratio of (theta/beta) has been used as an attention score [148, 149]. In fact, this ratio quantifies the attention level of the subject which somehow identifies the state of the brain.

The role of α and θ band power on attention has been previously studied [147, 150, 151]. Studies have shown that higher pre-stimulus α power represents low attention state. While, higher θ power represents higher attention state. In other words, the effect of alpha and theta band are related in an opposite way [71], which means that better performance is achieved by increasing theta band power and decreasing alpha band power. Hence, we define a novel coefficient which is computed from power of pre-cue EEG data over three frequency bands of theta, beta, and alpha.

Before calculating spectral powers, the recorded EEG signals are visually inspected and those trials with excessive eye-blinks are removed. The cleaned EEG data from all the channels are then filtered over θ (4-8 Hz), α (8-13 Hz), and β (16-24 Hz) frequency bands. In the next step, the filtered data are spatially filtered by means of Local Average Reference (LAR). The average activity of the closest neighboring electrodes is subtracted from each individual electrode.

The EEG band power in each of the three specified frequency bands is

calculated and normalized over all trials (Eq. 3.1):

$$\overline{\mathbf{E}}_i^2 = \frac{\mathbf{E}_i^2}{\sum_{j=1}^{N_T} \mathbf{E}_j^2} \tag{3.1}$$

where $\mathbf{E}_i \in \mathbb{R}^{c\times\tau}$ denotes the filtered single trial EEG measurement of the i^{th} trial; $\overline{\mathbf{E}}_i \in \mathbb{R}^{c\times\tau}$ denotes the normalized spectral power of the i^{th} trial over all trials; and n_t shows the number of trials; τ is the number of EEG samples per channel; and c denotes the number of channels. Subsequently, the normalized band power $\overline{\mathbf{E}}_i$ is averaged over pre-cue time segment according to Eq. 3.2.

$$\mathbf{P}^b = \sum_{i=1}^{n_t} \sum_{t\in pre\,cue} (\overline{\mathbf{E}}_{i,t}^b)^2 \tag{3.2}$$

where $\overline{\mathbf{E}}_i^b \in \mathbb{R}^{c\times\tau}$ denotes the filtered single trial EEG measurement of the i^{th} trial in frequency band of $b\in \{\theta, \alpha, \beta\}$; $\mathbf{P}^b \in \mathbb{R}^c$ denotes the average pre-cue frequency band-power over all trials, and τ denotes the number of EEG samples per channel. The proposed coefficient is finally defined from the averaged frequency band powers as follows:

$$F = \frac{\sum_{c\in C_\theta} p_c^\theta}{\sum_{c\in C_\alpha} p_c^\alpha + \sum_{c\in C_\beta} p_c^\beta} \tag{3.3}$$

where p_c^b denotes the average of pre-cue frequency band-power of c^{th} channel; b denotes the frequency band of $\in \{\theta, \alpha, \beta\}$; and C_θ, C_α, and C_β denote selected channels from frontal, parietal and central area.

Different regions of the brain have different activation patterns over specific frequency band. This led us to consider topographical information in our proposed coefficient. It can be assumed that a coefficient which is based on EEG rhythms computed from different brain regions might be more infor-

mative in predicting the performance of the users. It was shown previously in some studies that theta band power over frontal area and alpha band power over parietal area can somehow represent the attention of the subjects [147, 152]. Therefore, to have a more meaningful neurophysiological predictor, we calculate the pre-cue EEG band powers over different regions of the brain. Accordingly, theta, alpha and beta band powers in Eq. 3.3 are averaged over frontal area C_θ ={F3, Fz, F4}, parietal area C_α ={P7, P3, Pz, P4, P8} and central midline area C_β ={Cz, Cpz}.

To improve the proposed predictor in Eq. 3.3, some weights are assigned to different band powers. Therefore, Eq. 3.3 is modified by normalizing theta band power over weighted sum of alpha and beta band powers as follows:

$$F_{new} = \frac{\sum\limits_{c \in C_\theta} p_c^\theta}{\lambda \sum\limits_{c \in C_\alpha} p_c^\alpha + (1-\lambda) \sum\limits_{c \in C_\beta} p_c^\beta} \tag{3.4}$$

where $\lambda \in [0, 1]$ is a weighting factor.

3.4 Experimental setup

The EEG data was collected from 17 healthy subjects. Two out of 17 subjects were left handed. All the subjects participated in two different sessions: a calibration session on a first day and a non-feedback session on a separate day. No feedback was provided for the subjects during the experiment.

During the experiment a visual cue was displayed on the computer screen which informed the subject to perform either MI or idle. During MI trials, subjects were instructed to perform kinaesthetic hand MI due to their handedness. To define the idle state for subjects, they were instructed to perform mental counting in idle trials. The main reason was to make the idle state more consistent to reduce both inter- and intra-subject inconsistency during

the idle state. Prior to the experiments, subjects were instructed to minimize any physical movement and eye blinking throughout the EEG recording process. All subjects were asked for ethics approval and informed consent.

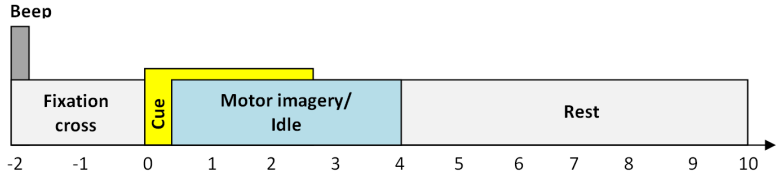


Figure 3.1: Timing scheme of each trial of experiment. A beep sound followed by a fixation cross on the screen notifies the subject about the start of each trial. A cue is shown on the screen at time 0. A subject starts performing either MI or idle right after the cue.

In the calibration session two runs of EEG data were collected. Each run comprised of 40 trials of MI and 40 trials of idle state and lasted about 16 minutes. Figure 3.1 represents the timing scheme of each trial. As shown, each trial comprised a preparatory segment of 2 s, the presentation of the visual cue for 4 s, and a rest segment of at least 6 s. Each trial lasted approximately 10 s, and a break period of at least 2 min was given after each run of EEG recording. The EEG data collected in calibration session was then used to calibrate a subject-specific model. During the non-feedback session three runs of EEG data were collected while subjects were performing MI of the chosen hand versus idle state. These three runs were almost similar to that of the calibration session, each lasted approximately 16 minutes and comprised of 40 trials of MI and 40 trials of idle state.

The Nuamps EEG acquisition hardware (http:\\www.nueroscan.com) with unipolar Ag/AgCl electrodes channels was used to collect EEG data. The recorded signal was digitally sampled at 250 Hz with a resolution of 22 bits for voltage ranges of ± 130 mV. EEG recordings from all 27 channels were band pass filtered from 0.05 to 40 Hz by the acquisition hardware.

3.5 Methods

3.5.1 Evaluating the classification performance of MI-BCI using FBCSP

The performances of subjects are evaluated based on the classification accuracy of the non-feedback session. As mentioned earlier in Section 3.4, a model is trained based on the EEG data collected in calibration session to detect MI in the non-feedback session. For estimating the performance of the users, FBCSP algorithm which was proposed in [117] is used. FBCSP selects subject-specific frequency bands and results in better performance in comparison to CSP. It has four stages:

- *Multiple frequency band-pass filtering*: A total of 9 Chebyshev Type II band-pass filters are used, namely, 4-8 Hz, 8-12 Hz, ... , 36-40 Hz.
- *Spatial filtering*: The CSP algorithm (please see Section 6.3.1) is applied to spatially filter the signal. The m first and last CSP filters are only selected for computing the features for each band.
- *Features selection*: The k best pair of features are selected among all 2m×9 features by using the mutual information-based best individual feature (MIBIF) algorithm.
- *Classification*: SVM classification algorithm is employed to model and classify the selected FBCSP features.

3.5.2 Proposed methodology for correlation analysis

We hypothesize that there is a positive correlation between the proposed EEG rhythm-based neurophysiological coefficient (Eq. 3.3) and subjects performance. To investigate our hypothesis a group level analysis is performed and Pearson's correlation coefficient between the proposed coefficient

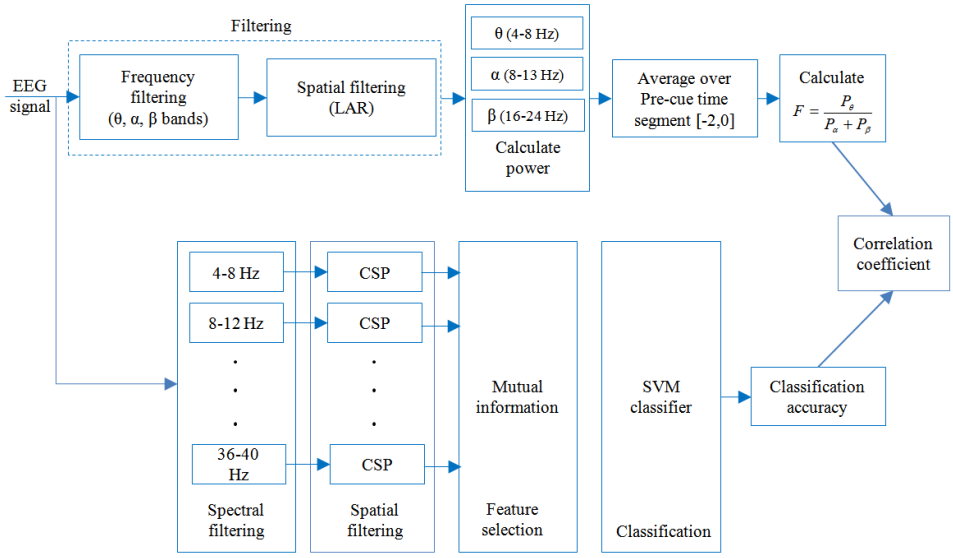


Figure 3.2: Methodology for calculating the correlation between the proposed coefficient and the classification performance of the subjects. The performance is calculated based on FBCSP method. The pre-cue time segment is selected according to the design of the experiment.

of all users and their accuracies is computed. The significance level of the test proves the validity of our initial assumption and shows how well the proposed novel coefficient can predict the performance of users. Figure 3.2 demonstrates our methodology for calculating the correlation between classification performance and the proposed coefficient.

3.6 Results

The recorded EEG signals were visually inspected and those trials with amplitude higher than 220 μvolts were rejected. This was mainly done to avoid theta band power being obscured by eye-blinks components in our proposed coefficient. Figure 3.3 shows the number of removed trials for each individual subject. As shown, for all of the subjects except one, less than 5% of trials are removed.

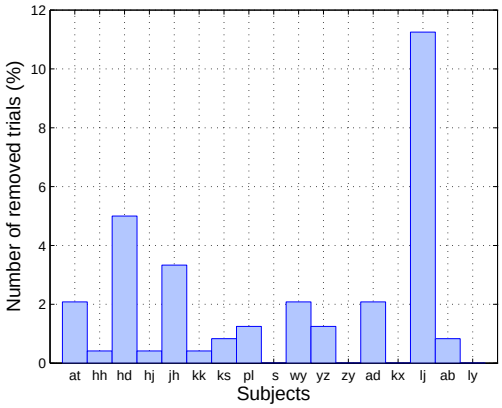


Figure 3.3: Number of trials (%) with excessive eye-blinks which were removed from our study.

3.6.1 Time course of EEG relative power

The proposed coefficient (Eq. 3.3) was computed from 2 s of pre-cue EEG data. In order to show the changes of EEG power over each of the studied frequency bands, the ERD/ERS time courses for each single channel were calculated based on the method described in [78].

Figure 3.4 shows ERD/ERS time courses of several channels from different regions of the brain. The ERD/ERS shown in this figure were calculated for subject *kk* who had high performance. The average relative power of each theta, alpha, and beta frequency bands over selected channels from frontal, parietal and central area are also plotted in Figure 3.5. As can be seen, EEG band powers start to decrease before a cue is provided for the subject at time 0. Hence, we may infer that pre-cue time segment contains information about the task that the subject is instructed to perform after cue timing.

3.6.2 Classification performance of the subjects

In this study, the accuracy of subjects is used as a measure of performance. Table 3.1 summarizes the performance accuracy of the the 17 healthy subjects. The session-to-session transfer accuracy showed the performance of

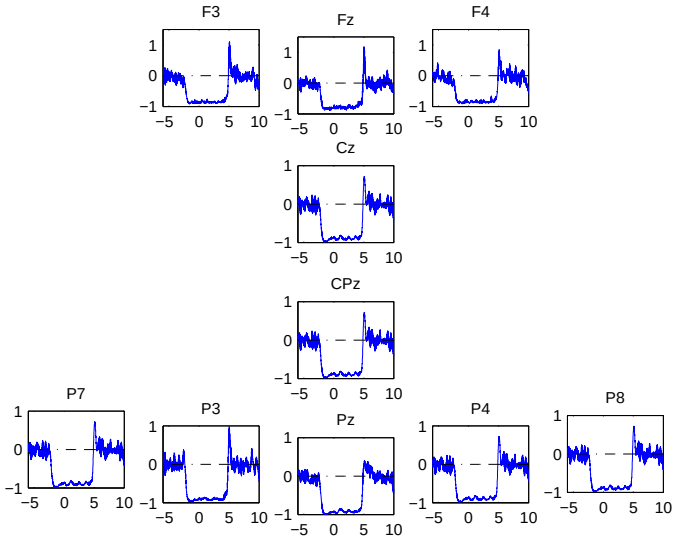


Figure 3.4: ERD/ERS time courses of healthy subject *kk* in different channels. The Frontal channels (F3, Fz, F4) show the theta activity. The central channels (Cz, CPz) show beta activity, and parietal channels (P7, P3, Pz, P4, P8) reveal alpha activity.

the users in evaluation session. As can be seen, the average performance of the subjects was 61.25%, and 10 out of 17 subjects had accuracies above 70% which is an acceptable threshold in most BCI systems. In other words, subjects with accuracy less than 70% were not successful in using BCI system and may be considered as subjects with BCI deficiency [153].

However, the session-to-session transfer performance can be affected by non-stationarity of EEG signal. Therefore, the 10×10-fold Cross-Validation (CV) accuracies of the subjects in both calibration session and evaluation session were calculated. The average CV accuracy of the calibration and evaluation session were 74.66% and 72.64%, respectively. The overall CV accuracies of evaluation session was higher than the session-to-session transfer results. The non-stationarity between calibration and evaluation session can be one of the possible reasons of such a low performance. Therefore, in our analysis we used 10×10-fold CV accuracies of the evaluation session as

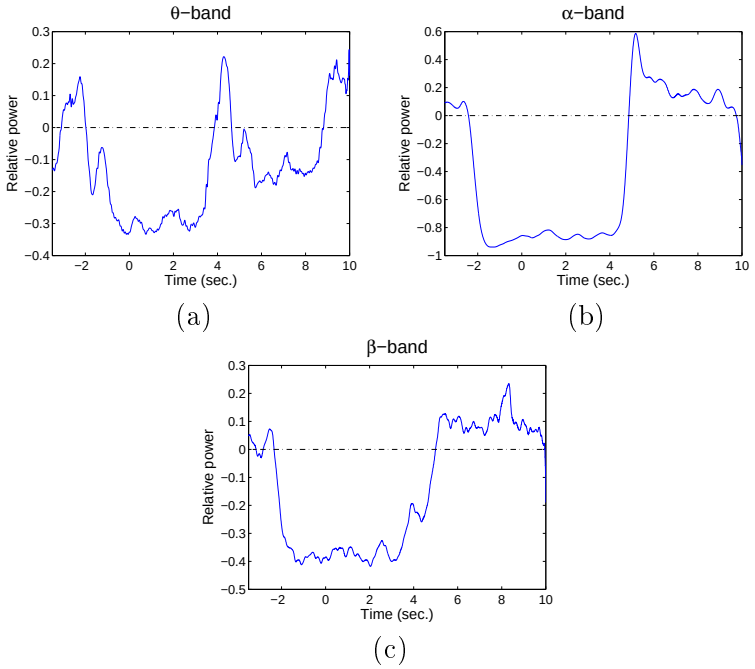


Figure 3.5: ERD/ERS time courses of healthy subject *kk*. Plots show (a) θ, (b) α, and (c) β bands ERD/ERS averaged over selected channels from frontal (F3, Fz, F4), central (Cz, Cpz) and parietal (P7, P3, Pz, P4, P8) area, respectively. The graphs are smoothed by means of moving average.

a measure of subjects' performance, since CV results are not affected by the inter-session non-stationarity. This can help us to better identify the BCI deficient subjects.

Figure 3.6 shows box plot of 10×10-fold CV accuracies for 17 healthy subjects in the non-feedback session. As shown, the median accuracy of the subjects varies from 50% to 95.8%. Subjects with accuracy less than 70% were considered as low performance subjects [136]. Hence, the subjects were divided into two groups: subjects with low performance (median=58.33%) and subjects with high performance (median=85.42%).

3.6.3 Correlation analysis

After calculating the proposed neurophysiological coefficient for each individual subject, a Pearson's correlation coefficient was then computed to as-

Table 3.1: Classification performance of 17 healthy subjects. The table includes the 10×10-fold cross-validation accuracies of calibration and non-feedback sessions, and also session-to-session transfer accuracies.

Subjects	10x10-CV acc (%) on calibration session	10x10-CV acc (%) on evaluation session	Session-to-session transfer acc (%)
at	59.13	64.08	53.75
hh	55.25	58.46	52.50
hd	64.50	54.63	63.33
hj	91.50	90.29	71.25
jh	79.83	74.71	70.00
kk	88.25	83.75	75.83
ks	62.42	85.29	77.08
pl	94.58	94.67	92.92
s	88.94	80.96	77.08
wy	48.17	51.38	57.50
yz	50.50	66.75	47.08
zy	87.33	81.17	72.92
ad	64.50	64.75	50.00
kx	81.25	62.21	86.67
lj	87.75	84.88	67.92
ab	92.25	79.46	82.92
ly	73.00	57.46	78.33
AVG	**74.66**	**72.64**	**61.25**

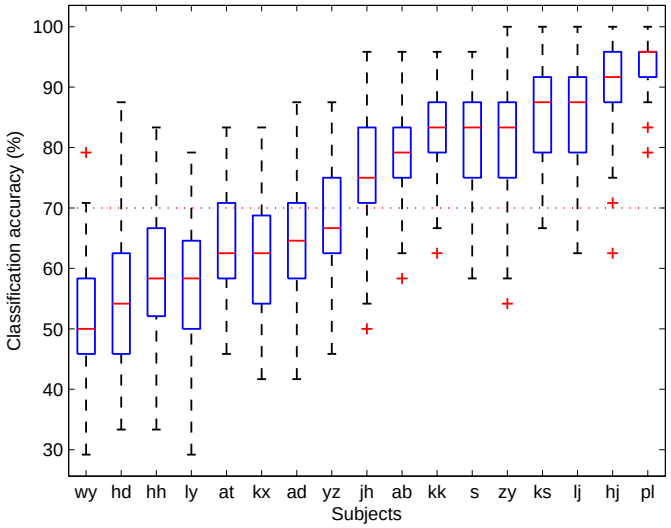


Figure 3.6: Box-plot of 10×10-fold cross-validation accuracy of 17 healthy subjects during non-feedback session.

sess the relationship between the proposed predictor and the classification performance of the subjects. The higher correlation value r and lower cor-

responding significant level p, demonstrate the strength of this correlation.

A significant positive correlation between the proposed predictor and CV accuracies (r=0.53, p=0.02) was achieved. This means that the proposed predictor explained as much as r^2=29% of the variance in classification performance of the subjects. Therefore, the correlation result indicates that subjects with higher classification accuracy have higher value of the predictor and vice versa. Figure 4.1 represents the values of the proposed predictor versus CV accuracies for each individual subject. As can be seen in this figure, high performance subjects (pl, s, ks, lj, kk, hj, ab, zy) have higher values of the predictor in comparison with low performance subjects (hh, hd, wy, kx, ly, at).

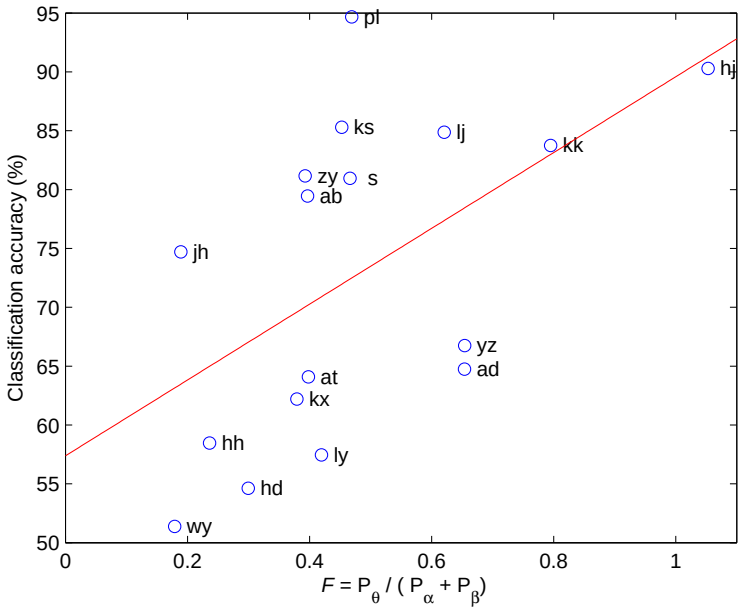


Figure 3.7: Correlation of the proposed EEG rhythm-based coefficient F with BCI classification accuracy. The accuracies are 10×10-fold CV accuracies of the users over non-feedback session. Each circle represents a healthy subject. The solid line (slope=0.32) is linear regression result.

The correlation between the proposed coefficient and the classification accuracy of the subjects without removing eye blinks was r=0.49, p=0.04.

This shows that removing the eye blinks enhances the correlation results.

3.6.4 Comparison of high and low performance subjects

A Mann-Whitney U-test was conducted to compare the value of the proposed predictor for low (Group 1) and high performance subjects (Group 2). The result of the test is shown in Figure 3.8.

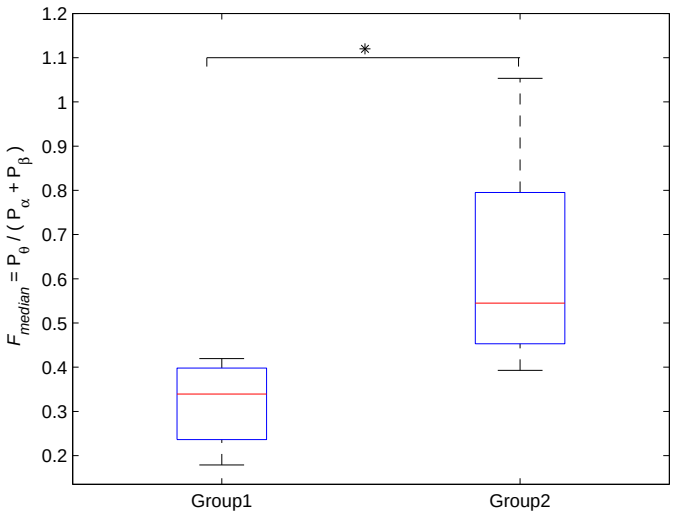


Figure 3.8: Comparing the proposed EEG rhythm-based coefficient F for two groups of subjects. Group 1 are six subjects with accuracy less than $35th$ percentile and Group 2 are the six subjects with accuracy above $65th$ percentile (Mann-Whitney U-test *p<0.05).

The subjects in Group 1 and Group 2 were chosen based on their classification performance. Group 1 contains subjects with accuracy less than $35th$ percentile, while Group 2 contains subjects with accuracy above $65th$ percentile. The results of the test showed that the value of the predictor for subjects with high performance (F_{median}=0.54) was significantly (p=0.008) higher than the value of the predictor for subjects with low performance (F_{median}=0.33). Therefore, it can be concluded that subjects with higher classification accuracy have significantly higher value of the proposed pre-

dictor or higher attention level. It should be noted that, the significance level of the result may change by grouping the subjects in other ways.

The values of the proposed predictor for two of the subjects *yz* and *ad* with low performance (accuracy less than 70%) are quite comparable to that of the subjects with high performance such as *kk*. On the other hand, subject *jh* with CV accuracy=74.7% has the lowest value of the predictor. As shown in Figure 3.6, these three subjects have moderate accuracies, that are placed somehow in the middle of the graph between two groups of subjects with high and low performance accuracies.

3.6.5 Weighting the proposed coefficient

The effect of weighting the coefficient is studied in Figure 3.9. This figure shows the results of group level correlation analysis for different values of λ in Eq. 3.4. The Pearson's correlation coefficient *r-value* and its corresponding *p-value* for each value of λ are plotted. As can be seen, for $\lambda \leq 0.67$ there was a significant correlation between F_{new} and performance. Highest correlation (r=0.63, p=0.007) was achieved by λ=0.16.

Figure 3.10 shows the group level correlation analysis for some of the selected λ values. As can be seen in this figure, for different values of weighting factor λ, the six lowest performance subjects had always significantly lower values of F_{new} in comparison with the six highest performance subjects.

3.7 Discussion

In this study, we proposed a novel neurophysiological coefficient to predict the performance of the subjects in EEG-based MI-BCI. The proposed predictor was defined based on pre-cue EEG. Several previous studies have shown that the ongoing oscillatory activity preceding an event has effect on sub-

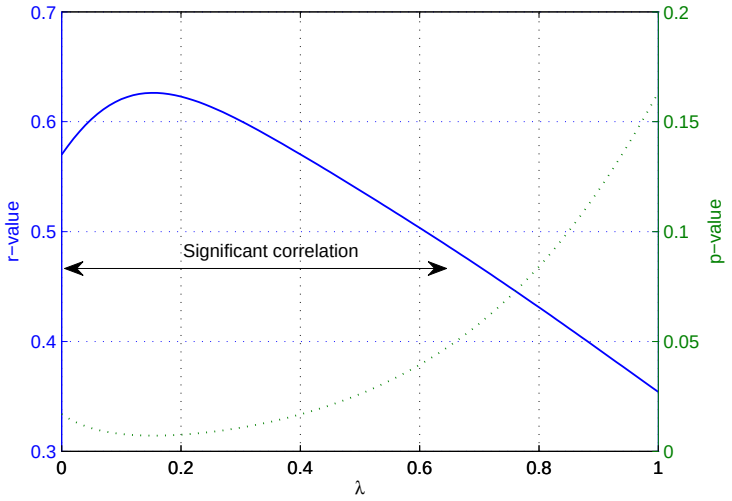


Figure 3.9: Analysis of the effect of weighting factor λ=[0 1] on the correlation coefficient between the proposed weighted coefficient F_{new} and CV accuracies of the subjects. The r-values (solid line) and p-values (dashed line) are the Pearson's correlation coefficients and their corresponding significant levels, respectively.

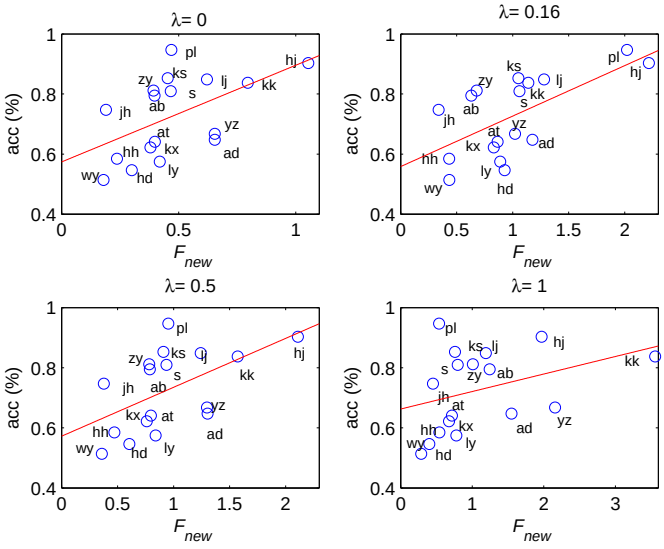


Figure 3.10: Correlation of the proposed weighted EEG rhythm-based coefficient F_{new} with 10×10-fold CV accuracies of the users over evaluation session for selected values of λ=0, 0.16, 0.5, and 1. Each circle represents a healthy subject. The solid lines are linear regression results.

sequent task outcome. As an example, pre-cue EEG data has been used to predict the performance of the subject in memory, mental workload and visual attention tasks. Therefore, it may infer that the state of the brain or readiness of the brain over the pre-cue time segment can predict the performance of the subject in performing motor imagery tasks. According to our experimental design, the subject was instructed to avoid any movement and be ready for the following task after hearing a beep sound. Before the beep sound there was a resting time during which the subject was allowed to be relaxed without any special instruction. Hence, the recorded signal contained several artifacts and it was not reliable to be used for our analysis. The time between hearing the beep sound and the cue timing was around 2 s.

As stated earlier the power of EEG signal over pre-cue time segment was computed for α, β, and θ frequency bands and used to quantify the current state of brain. ERD/ERS time course shown in Figure 3.5 reveals that the relative powers start to decrease 2 s before cue timing. Therefore, this time segment was used in calculating the attention level of the user prior to the start of trial at time 0.

Although the brain's functionalities during different tasks are not the same, it is common in all tasks that state of brain can affect the subject's performance. To define the brain's state we tried to capture spatial and spectral information in our proposed coefficient. Several studies proposed performance predictors which was reviewed in Section 3.2; however, pre-cue information of different frequency bands over different regions of the brain has not been used for prediction so far.

In the proposed coefficient theta band power is calculated over frontal area. Frontal theta activity has previously shown to be related to attention [147]. It has been shown that attentional processes and working memory are

closely related, which means that an increase in frontal theta activity is due to an increase of working memory load. In other words, higher attention level is expected when there is higher frontal theta activity. This can justify the reason of focusing on frontal theta as a part of the coefficient proposed for attention level quantification. According to our electrode setup the frontal channels F3, Fz, and F4 were selected to calculate the theta activity. Several studies have shown the role of pre-stimulus alpha over parietal and occipital area [142, 144, 145]. They concluded that lower alpha results in higher accuracy [144, 145] and faster reaction time [142]. Therefore, in our proposed coefficient, alpha band power was calculated over parietal area (P7, P3, Pz, P4, P8).

As we defined the brain's state by the attention level, we assumed that those subjects with higher accuracies had higher attention level, which means that they had higher values of the proposed neurophysiological coefficient. Due to the definition of our proposed coefficient in Eq. 3.3, higher attention level is achieved by having higher frontal theta and lower parietal alpha. The results of this chapter showed a significant positive correlation between the proposed coefficient and the accuracies of the users, which implies our initial assumption was correct. It has been previously investigated in [138] whether attention span as a psychological parameter could predict the BCI performance, nevertheless no significant results were achieved.

The group level analysis in this chapter also revealed that giving different weights to alpha and beta band power in Eq. 3.4 yields to higher correlation. This may suggest that EEG rhythms have different weights in quantifying the attention level of subjects. Although significant positive correlation was achieved in group level analysis, as shown in Figure 4.1, the value of the proposed coefficient was totally different for subject yz, ad, jh. These three subjects had moderate accuracies with different values of the

predictor. Hence, it may be assumed that the predicted accuracies for some of the subjects are not precise.

Currently our findings were tested on a group of 17 subjects with various range of performance. This can be considered as a limitation of our study. It would be appreciated to test the proposed coefficient on a bigger data set with larger number of subjects. In comparison with other previous predictors such as SMR predictor [35], this coefficient captures spatial information from the brain in addition to spectral information. In fact, the proposed coefficient is computed from the spectral power of pre-cue EEG data for specific rhythms over different regions of the brain. In contrast, the SMR predictor is computed from the alpha band over C3 and C4 channels, which are not the two channels with highest alpha activity for all the subjects. The proposed coefficient in this chapter cannot estimate the performance of the subjects from the resting state EEG data, this can be also considered as another limitation of the proposed method.

3.8 Summary

In this chapter, we demonstrated that pre-cue EEG rhythms contain useful information about the following MI task. It has been shown in several previous studies that pre-stimulus EEG data can be used for predicting the performance of the users for the following task regardless of the type of the task (i.e., memory task, mental workload, engagement, oddball). In fact, pre-cue EEG data can somehow show the current state of brain.

Defining the current state of brain is not straightforward, since it may be affected by different factors such as changes in subject's attention, concentration, engagement, mood and some other factors. Therefore, it can be concluded that by defining the current state of brain, we can predict the

performance of the subjects.

It was assumed that the attention level of the subjects can be an indicator of the brain's state. Accordingly, we tried to quantify the attention level. The proposed coefficient may be considered as one of the possible quantification of the attention level. In order to include topographic information in the proposed coefficient, power of EEG signal over different regions of the brain was computed. Therefore both spatial and spectral information were captured and thus the current state of brain was better represented.

The group level correlation analysis represented that the new proposed coefficient was positively correlated to the accuracies of users during MI versus idle task. The results suggested that subjects with higher (lower) accuracies have higher (lower) values of the proposed coefficient. However, from this study we cannot infer that this is a causal correlation.

In conclusion, although the results of this chapter were based on group level analysis, they are promising enough to lead us into a new experimental design that specifically helps subjects with BCI deficiency.

Chapter 4

Spatio-spectral based performance predictor

4.1 Introduction

MI induces changes in SMRs over sensorimotor areas [24]. The SMRs comprise oscillations of alpha (8-13 Hz) and beta (18-30 Hz) frequency bands. A decrease of SMR during MI task is known as ERD, while an increase of SMR is known as ERS [78]. Studies have shown that alpha ERD and beta ERS were observed during preparation, execution and imagination of motor activity over sensorimotor areas [24]. Hence, an EEG-based BCI system can be designed to decode the modulation of alpha and beta oscillations, but not all subjects can modulate their rhythms to control the BCI system properly. For some subjects BCI performance may drop below an acceptable threshold of 70% [34, 153].

Recently, several methods has been proposed to predict the performance of the MI-based BCI users [35, 38, 137–139]. Psychological predictors such as attention, personality or motivation have moderate but meaningful roles on BCI performance [138]. In contrary, neurophysiological predictors defined based on EEG rhythms has been shown to be more reliable in predicting the BCI performance than psychological predictors [35, 137] (please see Section 3.2 for more information about psychological and neurophysiological perfor-

mance predictors)

One of the proposed neurophysiological predictors [35] was based on the power spectral of two min EEG recordings of resting state (i.e., *relax with eyes open*). The average difference between the maximum peaks of the power spectral of the EEG and $1/f$ noise spectrum over two laplacian channels C3 and C4 was found to be significantly correlated (r=0.53) to the BCI performance of 80 healthy subjects. Therefore, it may be assumed that subjects with higher spectral peak at rest have higher BCI performance. The influence of the pre-stimulus subject-specific alpha band power on classification performance was also studied in [139]. They showed that the trials with higher pre-stimulus band power in comparison with trials with low pre-stimulus band power were better classified. However, only subjects with BCI performance above 70% who performed right versus left hand motor imagery were included in their study.

The above-mentioned neurophysiological predictors were all based on the power of EEG rhythms at predefined frequency bands. However, the recorded EEG is superposition of neural signals from several sources with a relatively strong background noise. Accordingly, Spatio-Spectral Decomposition (SSD) [154] was proposed to extract neural oscillations based on a linear decomposition of the recorded multichannel EEG. SSD maximises the EEG power at desired frequency band while minimizing the power at neighboring frequency bands, therefore it can extracts the neural oscillations in presence of strong background noise. It may be assumed that using SSD may yield a more reliable estimation of EEG activity over the frequency band of interest from multichannel EEG data. However, the SSD method had only been used only for analyzing brain sources.

In this chapter, we applied SSD to enhance the extracted alpha band activity that was shown previously to be correlated to the performance of

the MI-BCI system [24, 78]. On the other hand, it has been shown [35, 139] that alpha band power from pre-stimulus or resting state can predict the BCI performance of the users. Therefore, in this chapter we investigated whether the extracted SSD components with enhanced alpha band spectral characteristics can better predict the BCI performance of the users to mitigate the effect of strong background noise.

4.2 Spatio-spectral decomposition (SSD)

Spatio-spectral decomposition [154] has been proposed to extract neural oscillations form multichannel EEG signal. Compared to methods based on independent component analysis, SSD is a much faster method. SSD method finds spatial filters which maximize the signal to noise ratio (SNR). It concomitantly maximizes the power of EEG signal over a desired frequency band and minimizes the EEG power over neighboring frequency bands. Hence, the extracted components have peaky spectral profiles. SSD method can successfully extract oscillatory signals even in presence of low signal to noise ratio [154]. In the following paragraphs, the main computational steps of SSD are briefly explained.

Let $\mathbf{E}_s \in \mathbb{R}^{\tau \times c}$ and $\mathbf{E}_n \in \mathbb{R}^{\tau \times c}$ be the filtered EEG signal over desired frequency band f (e.g., 8-13 Hz) and neighboring frequency bands $f + \Delta f$ and $f - \Delta f$ (e.g., 6-8 and 13-15 Hz), where τ and c are the number of samples and channels, respectively. Therefore, the covariance matrices of signal ($\mathbf{C}_s$) and noise ($\mathbf{C}_n$) are estimated as follows:

$$\mathbf{C}_s = \frac{\mathbf{E}_s^T \mathbf{E}_s}{\tau} \tag{4.1}$$

$$\mathbf{C}_n = \frac{\mathbf{E}_n^T \mathbf{E}_n}{\tau} \tag{4.2}$$

The spatial filters ($\mathbf{w}_i \in \mathbb{R}^{c\times 1}$) are then determined by maximizing the following SNR [154]:

$$SNR(\mathbf{w}_i) = \frac{\mathbf{w}_i^T \mathbf{C}_s \mathbf{w}_i}{\mathbf{w}_i^T \mathbf{C}_n \mathbf{w}_i} \tag{4.3}$$

which leads to generalized eigenvalue decomposition [155]:

$$\mathbf{C}_s \mathbf{w}_i = \lambda_i \mathbf{C}_n \mathbf{w}_i \tag{4.4}$$

where λ_i is the corresponding eigenvalue. The extracted SSD components $\widetilde{\mathbf{X}}$ can be obtained by:

$$\widetilde{\mathbf{X}} = \mathbf{X}\mathbf{W} \tag{4.5}$$

where $\mathbf{X} \in \mathbb{R}^{\tau\times c}$ denotes the recorded EEG signal before band-pass filtering, $\widetilde{\mathbf{X}} = [\widetilde{\mathbf{x}}(1), \widetilde{\mathbf{x}}(2), \cdots, \widetilde{\mathbf{x}}(\tau)]^T$, and $\widetilde{\mathbf{x}}(t) \in \mathbb{R}^{1\times n_m}$; n_m denotes the selected number of SSD filters used for spatially filtering the EEG data ($n_m \leq c$); c denotes the total number of channels; and $\mathbf{W} \in \mathbb{R}^{c\times c}$ denotes a spatial filter where its columns $\mathbf{W} = (\mathbf{w}_1, \cdots, \mathbf{w}_{n_m})$ are sorted in descending order according to their corresponding SNR values.

4.3 SSD-based performance predictor

It has been shown in previous studies [35, 139] that higher SMR amplitude over resting state (*relax with eyes open*) and also over pre-cue EEG of feedback runs yielded higher classification performance. Therefore, extracting alpha activity is crucial part for predicting the BCI performance of the users. The SSD method can extract the components with highest SNR over a frequency band of interest. Therefore, we hypothesise that the power of SSD components with highest alpha activity are positively correlated with performance of the subjects and thus can be used for predicting the performance of the subjects. Our proposed SSD-based performance predictor in

this chapter is based on the power of SSD components. To define our predictor, the spatially filtered EEG data $\widetilde{\mathbf{X}}$ in Eq. 4.5 is band-pass filtered over desired frequency band f and the power of SSD components $\mathbf{p}_{\widetilde{X}}$ is estimated as follows:

$$\mathbf{p}_{\widetilde{X}}(t) = \|\widetilde{\mathbf{x}}(t)\|^2 \tag{4.6}$$

where $\mathbf{p}_{\widetilde{X}}(t) \in \mathbb{R}^{1\times n_m}$. Subsequently, the power of SSD components $\mathbf{p}_{\widetilde{X}}(t)$ is averaged over the selected number of components and time in Eq. 4.7.

$$\overline{p} = \frac{1}{n_m}\sum_{i=1}^{n_m}\frac{1}{\tau}\sum_{t=1}^{\tau} p_{\widetilde{X}}(t,i) \tag{4.7}$$

where $\overline{p}$ can be named as SSD-based performance predictor.

4.4 Experimental data

In this study, a large data set of 80 healthy BCI-novice subjects is used [35]. Each subject participated in one session that comprised of calibration runs and feedback runs recorded on the same day.

The subjects sat on a comfortable chair with their arms being relaxed on an armrest during the whole experiment. The EEG signals were recorded from 119 Ag/AgCl electrodes using multi-channel EEG amplifier (BrainAmp DC by Brain Products, Munich, Germany). The recorded brain activity was sampled at 1000 Hz and band-pass filtered between 0.05 and 200 Hz. Electroencephalogram (EOG) and electromyogram (EMG) were also recorded.

At the start of the session and before the calibration session, ten periods of 15 s of alternating *relax with eyes open* and *relax with eyes closed* EEG were recorded. In calibration runs, no feedback was provided for the subjects. The subjects were instructed to perform right, left, or foot motor imagery task. A total of 225 trials of MI were recorded which comprised of 75 trials for

each MI task. The two classes which led to the highest classification accuracy were selected. During the feedback runs, the subjects were instructed to perform MI of the two selected classes. For a majority of the subjects, three runs of 100 trials with BCI feedback were recorded. For more information on the experimental set up, please refer to [35, 138].

4.5 Methods

4.5.1 Artifact removal

The EOG channels were used to remove the eye blinks from the recorded EEG signal. To minimize the eye motions and blinks, a method previously proposed in [156] was used. Let $\mathbf{X} \in \mathbb{R}^{c \times \tau}$ be the recorded EEG and $\mathbf{A}$ represents the coupling of sources $\hat{\mathbf{S}} \in \mathbb{R}^{c \times \tau}$ which are desired to be removed (i.e., eye blinks); c denotes the number of channels; τ denotes the number of samples in each channel. The sources can be estimated from EEG as, $\hat{\mathbf{S}} = \mathbf{A}^{\dagger}\mathbf{X}$, where $\mathbf{A}^{\dagger}$ denotes the pseudo inverse of $\mathbf{A}$. The sensor activity generated by the undesired sources is $\mathbf{X}_{\parallel} = \mathbf{A}\hat{\mathbf{S}}$. Therefore, this activity can be subtracted from original EEG as follows:

$$\mathbf{X}_{\perp} = \mathbf{X} - \mathbf{X}_{\parallel} = (\mathbf{I} - \mathbf{A}\mathbf{A}^{\dagger})\mathbf{X} \tag{4.8}$$

where $\mathbf{X}_{\perp}$ is the EEG data without eye blinks.

4.5.2 Evaluating BCI performance of the subjects

For evaluating the BCI performance of the subjects, the recorded calibration runs were filtered in a subject-specific frequency band and time interval [88]. On the next step, CSP was used to spatially filter the data. The number of CSP filters were selected heuristically. Finally, the log-variance of the filtered

data were calculated and used as features for training the LDA classifier.

The performance of the subjects during the online feedback runs were evaluated based on the adapted LDA classifier. The first initial 20 trials of each feedback run were used for bias adaptation of the LDA classifier. The final reported performance is the average performance of the recorded runs. More information can be found in [35].

4.5.3 Correlation analysis

In this chapter, we defined the proposed predictor over resting state and pre-cue time segment. Therefore, the SSD-based predictor for both cases were computed using Eq. 4.5 to 4.7. Subsequently, the Pearson's correlation coefficient between the derived predictor and the feedback BCI performance of the subjects was calculated.

4.6 Results

The SSD-based predictor proposed in the chapter was used to predict the BCI performance of the subjects. Its effectiveness was evaluated on resting state (*relax with eyes open*) EEG, and pre-cue EEG from the feedback runs and was compared to the previously proposed spectral-based predictor in [35] (see Section 4.7).

4.6.1 BCI performance results

Figure 1 summarizes the accuracies of all 80 subjects during feedback runs. The classification performances vary considerably, ranging from 47.8% to 100% with a mean of 74.4% $\pm$ 16.5%. The intra-subject variability between runs ranged from 0% to 19.8%. 30 subjects performed right versus left hand MI during the feedback runs (Group 1, mean=81.13%), 34 subjects

performed left hand versus foot MI (Group 2, mean=71.74%), and the resting 16 subjects performed foot versus right hand motor imagery. Group 3 has the lowest average performance (Group 3, mean=67.52%) comparing to the other two groups.

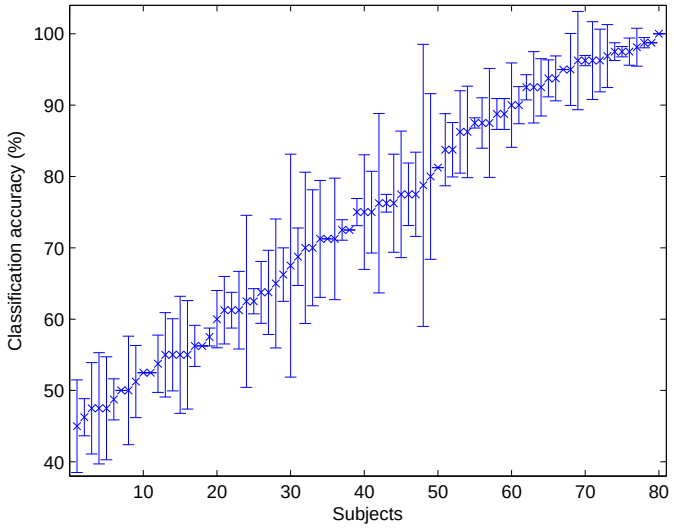


Figure 4.1: Classification accuracies of the feedback runs. Each cross (×) represents the median of classification accuracies of the recorded runs for each subject, and the vertical lines represent the corresponding standard deviations. Subjects are sorted according to their BCI performance.

4.6.2 Resting state EEG analysis

As stated earlier in Section 4.4, during the resting state 10 trials of *relax with eyes open* and 10 trials of *relax with eyes closed* were recorded. In this chapter we only analyze eyes open trials of resting state.

To choose the number of SSD components n_m, one may only select the first SSD component, since it has the highest SNR value and the most peaky profile. However, for some subjects using two or more components may enhance the results. In order to define a consistent measure for all the subjects, we performed a group level correlation analysis between the SSD-

based predictor and BCI performance of the subjects using different number of components n_m. The results are shown in Figure 4.2. Each dot represents the correlation result using the corresponding n_m components. The figure shows the effects of using the first 30 components $n_m = \{1, 2, \cdots, 30\}$. Group level analysis of all subjects showed the maximum significant correlation of $r = 0.53$, when the first two SSD components were used. However, by using more than two SSD components the correlation started to decrease. Moreover, the highest correlation for Group 1 (RL) and Group 2 (LF) of participants were $r = 0.57$ and $r = 0.49$ when using the first two SSD components. However, Group 3 (FR) of subjects had the highest correlation $r = 0.73$ using only the first SSD component. In summary, due to group level analysis we selected two SSD components.

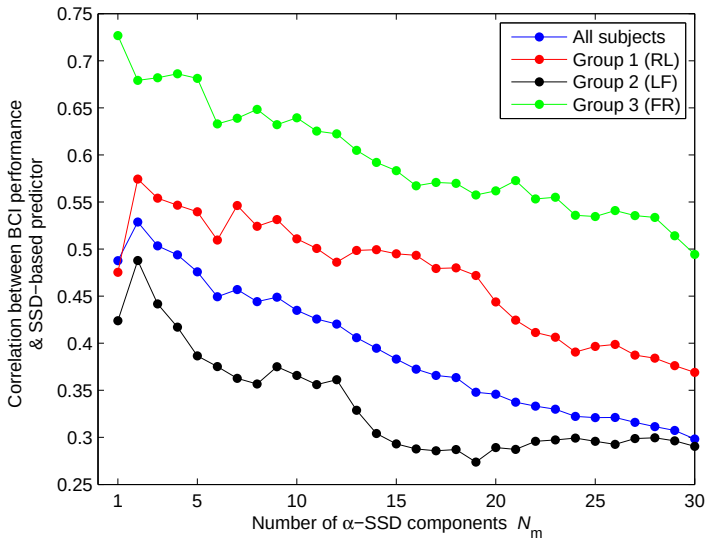


Figure 4.2: The effect of using different number of SSD components on the correlation (Pearson) between the SSD-based predictor and BCI performance of the subjects. Alpha-SSD components are extracted from resting state (*relax with eyes-open*). The blue dotted line represents the results of group analysis using all subjects. The red, black, and green dotted lines are the within group results for Group 1 (RL), Group 2 (LF), and Group 3 (FR), respectively. Each dot (•) represents a correlation result when using n_m SSD components.

Analyzing the results of SSD-based predictor suggested that subjects

with higher performance had higher value of the proposed predictor. As stated previously in [34], 48 subjects (Cat.I) had accuracy above 70% in both calibration and feedback runs, 14 subjects (Cat.II) had feedback accuracy less than 70%, and the resting 18 subjects (Cat.III) had performance less than 70% during the calibration session. Analysis of Variance (ANOVA) showed that the SSD-based predictor was significantly different between these three categories of users $F(2,77)=10.778$, $p<0.01$. Post hoc comparisons using the Bonferroni test revealed that Cat.I of has significantly higher value of SSD-based predictor than Cat. II ($p=0.02$) and Cat.III ($p<0.01$). This is similar to previous findings that subjects with higher alpha activity at rest state have higher BCI performance [35].

The selected components for one of the subjects with high feedback performance (98.12%) are visualized in Figure 4.3. This subject performed right versus left hand motor imagery during the feedback session. The first/second row of the figure presents the first/second SSD pattern, filter, and spectra of resting state. As expected, the spectra of the first extracted SSD component in comparison with the second one has higher peak at alpha band.

As can be seen in Figure 4.3, both components are extracted from the sources mostly located over sensorimotor area. However, the location of the sources are not the same for all the subjects. Figure 4.4 shows the first extracted α-SSD filter for six selected subjects. As shown, although higher weights are assigned to the channels over sensorimotor area, the channels are not exactly the same for different subjects. This shows that highest alpha activity is not always over C3 and C4 channels.

4.6.3 Pre-cue EEG analysis

Here in this section, we investigate the feasibility of predicting the performance of the users based on the SSD-based predictor computed from pre-cue

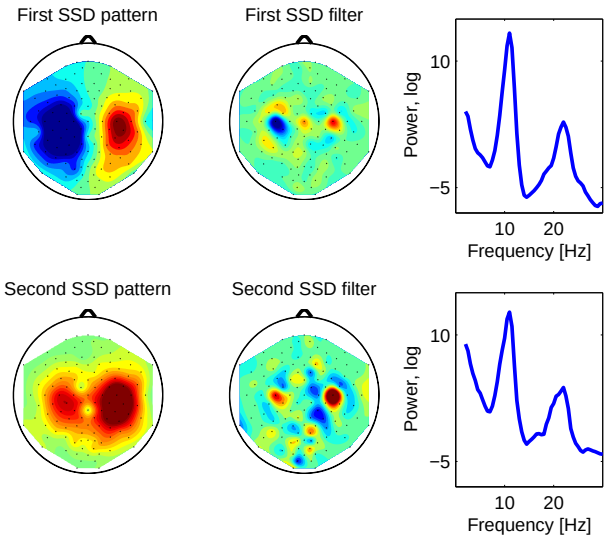


Figure 4.3: The first two α-SSD patterns, and their corresponding filters and spectras extracted from resting state (*relax eyes-open*) for subject 19 with feedback accuracy of 98.12%. The color-scale is in arbitrary units.

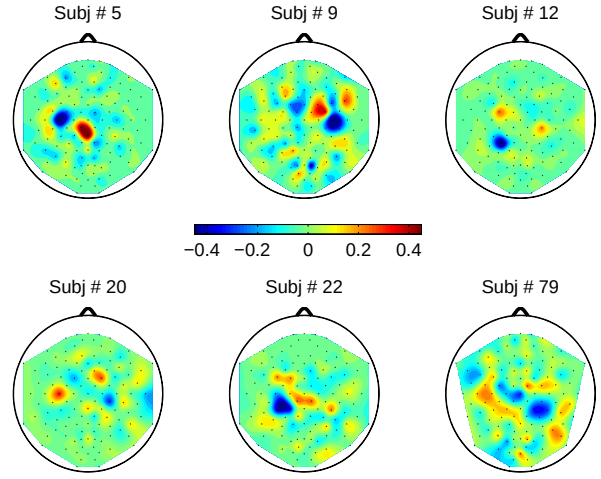


Figure 4.4: First SSD filter over the resting state (*relax eyes-open*) for the six selected subjects. The color-scale is in arbitrary units.

time segment of EEG. The SSD-based predictor (Eq. 4.7) was computed over 1 s of pre-cue EEG data. Then, the group level correlation between the SSD-based predictors and BCI performances was calculated. The best number of

SSD components were selected from the resting state analysis. It was shown that $n_m = 2$ results in the highest correlation between the proposed predictor and the BCI performance of the subjects (Section 4.6.3). Therefore, we used the first two SSD components computed from pre-cue EEG. Group level correlation analysis showed significant correlation $r = 0.39$ between SSD-based predictor and BCI performance of the subjects. Within group correlation analysis for Group 1 (RL) and Group 2 (LF) also showed significant correlation of $r = 0.40$ and $r = 0.34$. Therefore, it may be inferred that the SSD-based predictor computed from resting state EEG in comparison with the predictor computed from pre-cue EEG can better predict the performance of the subjects. However, the correlation result for Group 3 (FR) was much higher $r = 0.80$ even in comparison with resting state analysis. Although two components were selected for pre-cue analysis, we also investigated the effect of choosing various number of SSD components n_m on the correlation between the pre-cue SSD predictor and BCI performance of the subjects. Figure 4.5 shows the correlation results for different n_m values.

As shown, the within group analysis results of Group 1 and Group 2 are somehow similar to the results of group level analysis. Similar to resting state results in Section 4.6.3, within group analysis of Group 3 showed higher correlation in overall. Although the resting state analysis showed that the best number of components was $n_m = 2$, the highest correlation for Group 3 of subjects was $r = 0.83$ when $n_m = 5$. In conclusion, the results showed selecting the first initial SSD components of resting state or pre-cue EEG resulted in the highest correlation value. ANOVA test showed that the SSD-based predictor for Cat.I, II, and III of subjects was significantly different $F(2,77)=5.331$, $p=0.007$. Bonferroni test proved that the significant difference was only between Cat.I and Cat.III $p=0.007$. This finding shows that subject with higher predictor value for pre-cue EEG may result in better

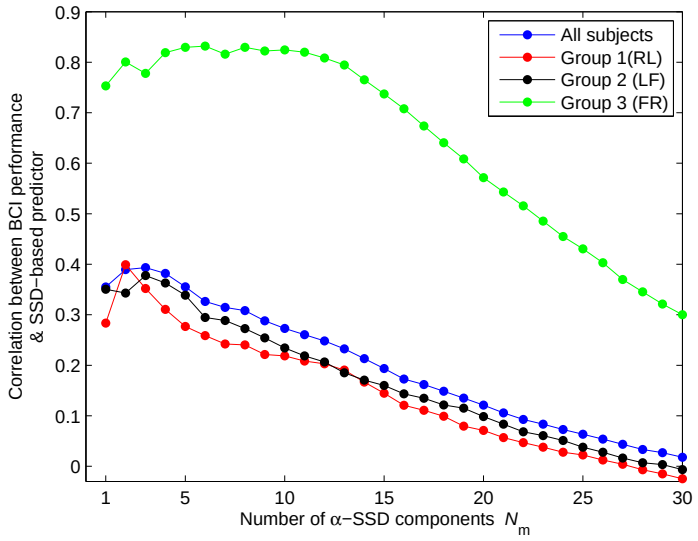


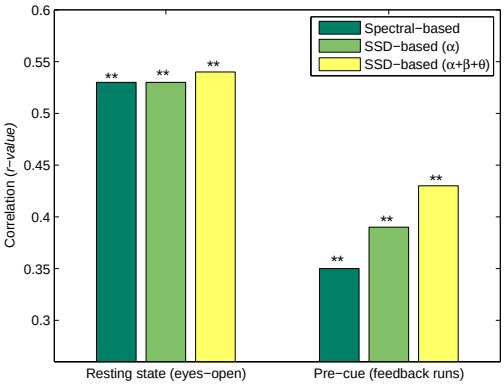
Figure 4.5: The effect of using different number of SSD components n_m on the correlation (Pearson) between the pre-cue SSD-based predictor and BCI performance of the subjects. The blue dotted line represents group analysis. The red, black, and green dotted lines are the within group results for Group 1 (RL), Group 2 (LF), and Group 3 (FR), respectively. Each dot (•) represents a correlation result when using n_m SSD components.

BCI performance.

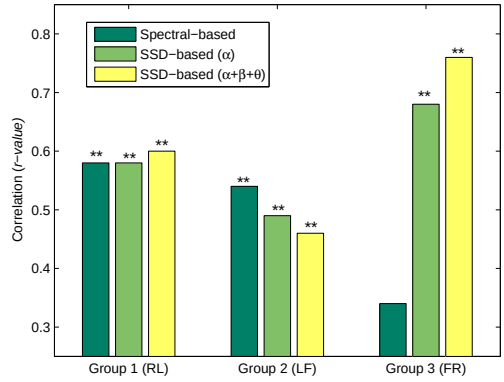
4.7 Comparison of the proposed SSD-based predictor with the spectral-based predictor

The results of the proposed SSD-based predictor are compared to the use of an existing spectral-based predictor or SMR predictor. The predictor proposed in [35] was calculated based on the power spectrum density (PSD) of the resting state EEG over two laplacian channels C3 and C4. The average difference of the estimated PSD curve and $1/f$ noise over two laplacian channels was defined as the SMR predictor.

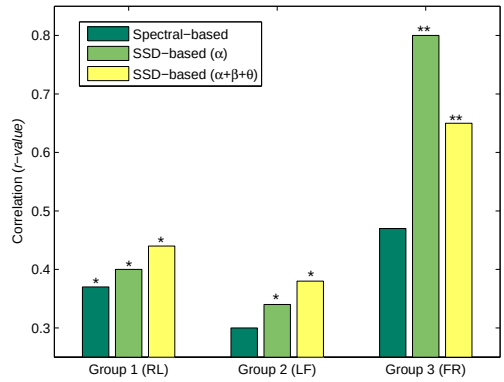
Figure 4.6 shows the group level correlation results of the proposed (SSD-based predictor) and the existing predictor (spectral-based predictor) for both resting state and pre-cue EEG. Two SSD components were selected for



(a)



(b)



(c)

Figure 4.6: Comparison of the proposed SSD-based predictor and the existing spectral-based predictor. The SSD-based predictor was computed using α components and also by incorporating other spectral components of β, and θ. (a) Group level (80 subjects) correlation results for two different conditions: resting state (*relax eyes-open*) and pre-cue of feedback runs; Correlation results of three different groups of subjects over (b) resting state and (c) pre-cue of feedback runs. (*p-value < 0.05, **p-value < 0.01)

computing the SSD-based predictor. The group level results of resting state for both methods were similar (r=0.53, p <0.01). However, the proposed SSD-based predictor outperformed the existing method over pre-cue EEG of feedback runs. The overall resting state correlation results were higher than pre-cue EEG correlation results. Therefore, it may be concluded that recording a short resting condition prior to the experiment is a good indicator of BCI users performance.

Figure 4.6 also demonstrates the within group analysis of Group 1 (RL), Group 2 (LF), and Group 3 (FR). As can be seen, within group analysis supports our results of group level analysis. Pre-cue EEG analysis of each group shows that the prediction role of the proposed predictor is not affected by the type of the motor imagery tasks.

4.8 Enhancing SSD-based predictor by incorporating other oscillatory activities

The role of alpha band power on motor activities has been shown in several studies [78, 93, 94, 139]. The proposed predictor in this chapter is based on the alpha-SSD components of EEG, computed from resting state or pre-cue time segment. However, spectral analysis of EEG during motor task has revealed that for some subjects the role of beta band is greater than alpha band [24]. Therefore, using beta band information may be also helpful in defining our predictor.

On the other hand, pre-stimulus theta [71, 141, 146] has been shown to affect the attention and memory performance [147]. It has also been proven that attention has a moderate but meaning role on performance of the users [138]. Therefore, we may assume that using theta and beta oscillatory activities may better predict the BCI performance of the users. So far, the

SSD-based predictor was computed based on the average of the first two alpha-SSD components. In order to test our hypothesis, the SSD-predictor is computed from average power of alpha, beta and theta components. As can be seen in Figure 4.6, defining the predictor based on alpha, beta, and theta components improves the group level correlation results of resting state $r = 0.54$ and pre-cue EEG of feedback runs $r = 0.43$.

4.9 Discussion

In this chapter, we proposed a novel neurophysiological predictor based on EEG rhythms with high SNR value. The proposed predictor used SSD to enhance extraction of the neural oscillatory activities over the frequency band of interest. Since alpha band power had been shown to be a reliable indicator for MI performance [35], the alpha-SSD components over resting state (*relax with eyes open*) and pre-cue EEG from the feedback runs were used to predict the performance of the users. The results showed that SSD components of alpha, beta, and theta from pre-cue EEG yielded better prediction. Although the role of theta oscillatory activity on motor tasks is not well studied, and there is insufficient evidence to show that the theta band activity has direct influence on the performance of a motor task, it may reflect the attention level of users as shown in [45]. Therefore, this may explain why subjects with higher alpha, beta and theta power performed better.

Although the results of SSD-based predictor over pre-cue time segment were significant ($r = 0.43$), they were not as good as the resting state results ($r = 0.54$). A study by Blankertz et al. has shown that the predictor computed from pre-cue EEG of feedback runs can better predict the BCI performance [139]. However, they only analyzed the data from subjects with accuracy above 70% who performed right versus left hand motor imagery.

However, in this chapter, all recorded trials and all subjects were analyzed and this could yield to different results in the two studies.

The spectral-based predictor proposed in [35] is computed over two laplacian channels C3 and C4. However, the foci of the minimum classification error may not be always C3 and C4. The foot motor imagery does not create a very focal cortical pattern [35]. Moreover, the best SMR channels are not the same for all the subjects [139]. By changing the laplacian channels from C3 and C4 to CP3 and CP4, the group level correlation result decreased from $r = 0.53$ to $r = 0.30$. This shows the dependency of the method on the selected SMR channels. Therefore, it is important to calculate the SMR predictor over subject specific channels. In contrast, SSD-based predictor can be applied on multi-channel EEG data to find the components with highest SNR over the frequency band of interest. Therefore, it is not necessary to select specific channels for subjects. As can be seen in Figure 4.6 (c) for the class combination of foot versus right hand motor imagery the correlation result of SSD-based predictor is much higher $r = 0.80$ than spectral-based predictor $r = 0.47$. Therefore, it may be assumed that the oscillatory activity can be better captured by SSD-based predictor.

The proposed SSD-based predictor and spectral-based predictor can both work properly in presence of noise. EEG signal is typically contaminated by different noise sources. SSD method extracts the neural oscillations from EEG signal even with a poor SNR [154]. Spectral-based predictor also discards the effect of noise by applying laplacian filtering and removing the noise floor values from the estimated PSD curves, As stated in [35], the correlation between SMR predictor and performance of the subjects decreases from $r = 0.53$ to $r = 0.33$ if the noise is not subtracted.

4.10 Summary

Despite several years of research, the reason for poor BCI performance of subjects is still unknown. To investigate this issue neurophysiological predictors have been proposed to investigate the variation in BCI performance among subjects. In this work, the SSD-based neurophysiological performance predictor is proposed to extract informative spatial-spectral components to enhance signal-to-noise ratio of the spectral characteristics. This will result in the better prediction of BCI performance.

The results showed that the proposed predictor outperforms the existing spectral-based predictor on both resting state *with eyes open* and pre-cue EEG of the feedback runs. This can be attributed to the enhanced extraction of oscillatory activities from cortical sources that resulted in a high correlation even for a group with a limited number of subjects (Group 3 with 16 subjects). Moreover, the pre-cue EEG results on different combination of MI tasks showed that the prediction results is not affected by the type of the MI task performed.

In conclusion, the proposed SSD-based method is an effective predictor of BCI performance. It can be used to better understand the brain topographies of specific frequency band over a desired time segment. Analyzing the SSD patterns also showed high performance subjects had higher resting state alpha over sensorimotor area. Therefore, designing a novel experiment which aimed to enhance EEG rhythms over specific regions of the brain can lead to improve BCI performance and address BCI deficiency problem..

Chapter 5

The Role of NeuroFeedback Training on Improving MI-BCI Performance

5.1 Introduction

Studies have shown that the MI-BCI performance of users varies extensively. The reason of why a considerable number of subjects are BCI deficient needs to be investigated. One of the possible reasons could be the subject inability in modulating the brain rhythms. Over the past years, some performance predictors have been proposed to assess the BCI performance of the users prior to the experiment. However, these predictors have not been used in real applications, no experiment has been designed so far to help those detected poor performance subjects control a BCI system better.

In Chapter 4 of this thesis, a novel neurophysiological performance predictor was proposed. The SSD-based performance predictor extracted the spectral components through performing spatio-spectral decomposition. The proposed predictor that was computed from the resting state EEG data was positively correlated to the BCI performance of the users. In addition, the results suggested that the role of alpha band on predicting the performance of the subjects was more significant comparing to theta and beta frequency

bands. Therefore, it may be assumed that enhancing the resting-state alpha rhythm results in improving the BCI performance and thus better control of a BCI system.

Neurofeedback training (NFT) is one of the possible methods used for enhancing the brain rhythms such as alpha rhythm [157]. Neurofeedback is one type of biofeedbacks which guides a subject to self-regulate his/her brain rhythms. During several training sessions, a subject may explore distinctive strategies to figure out how to regulate his brain rhythms. NFT has been utilized for different purposes [157–159]. As an example, NFT was a useful method for treatment of several neurological and psychiatric disorders such as ADHD [160].

NFT has been previously used for regulating SMRs (please see Section 5.2). In light of positive effects of neurofeedback on regulating the brain rhythms, it may be inferred that NFT can be used to enhance the resting state alpha rhythms. On the other hand, it was shown earlier in Chapter 4 that the resting-state α-SSD components were positively correlated to BCI performance of the subjects. Therefore, we hypothesize that enhancing the resting state α-SSD components through NFT sessions yields better BCI performance. However, to the best of our knowledge the impact of improving resting state SMR on the MI-BCI performance has not been studied so far. In this chapter, a novel experimental design is proposed to explore the impact of increasing resting state α-SSD components on the BCI performance of the users. Our proposed design is fully described in Section 5.3.

5.2 SMR regulation with NFT

SMR amplitude in resting state is dominant. However, planning, imagination or execution of movement desynchronizes the SMR amplitude [78]. Studies

have shown that subjects learn to self regulate their SMRs through several NFT sessions [67, 157–159, 161–164]. In these studies, subjects regulated their SMRs without any definite instruction and received a visual feedback corresponding to their SMR level.

SMR (12–15 Hz) was significantly enhanced by using an instrumental conditioning over several sessions of training [162]. A continuous visual feedback corresponding to the SMR amplitude was provided for the subjects to keep the SMR above a certain threshold for a specific period of time. A new training approach over C3-A1 and C4-A2 channels was proposed for alpha (8–12 Hz) training [67]. The relative alpha power in both eyes open and eyes-closed condition enhanced over training sessions. The protocol proposed in [67] also included a follow up session three months after finishing the training sessions.

A recent study [163] used neurofeedback for enhancing the alpha band. They showed the resting alpha rhythm prior the NFT sessions [163] was correlated to NFT performance. Both eyes open and eyes-closed resting state alpha were positively correlated to learning indices and could predict the learning ability over NFT sessions. However, the relative power over the training sessions did not show significant improvement.

In summary, SMRs enhancement through several neurofeedback training sessions is plausible. However, the neurofeedback training protocols of the above mentioned studies are not identical, which were defined based on the main objective of the studies. None of the previous studies examined the effects of neurofeedback on the BCI performance.

5.3 Study design

5.3.1 Participants

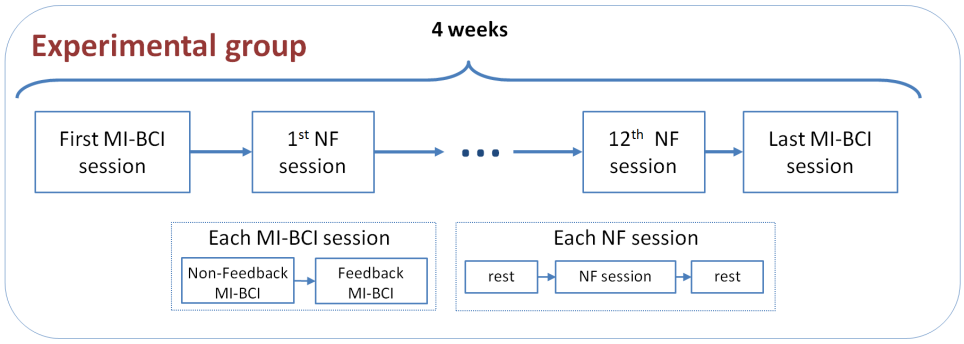
A total number of 13 healthy participants (6 female, 7 male) took part in this study after giving written informed consent. Participants were randomly assigned to two groups: an experimental group (3 males and 3 females, mean age 25.5 years), and a control group (3 males and 4 females, mean age 27.29 years). The mean age of two groups was not statistically different ($F(1,11)=0.953$, $p=0.35$). All participants except two subjects from the experimental group were novices for BCI experiment and all of them had not performed NF experiment before. Participants were only informed about the purpose of the study; however, they did not know about the grouping design.

The experimental design for the experimental and control group are shown in Figure 5.1 (a) and (b). Subjects of the experimental group participated in a MI-BCI session followed by 12 sessions of NF and a final MI-BCI session. The control group participated in only two MI-BCI sessions four weeks apart to diminish the learning effect.

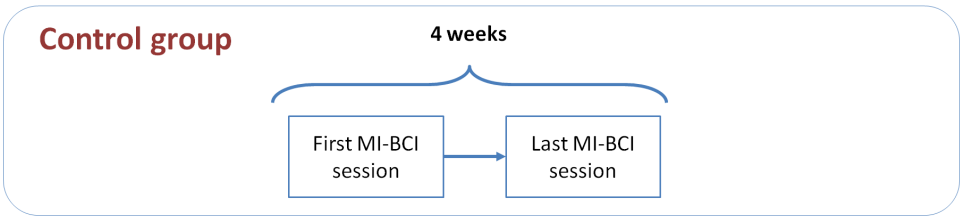
5.3.2 MI-BCI session

Each MI-BCI session started by collecting 5 minutes resting state EEG data, followed by 2 non-feedback calibration runs and 2 feedback runs. Within the resting state, a visual cue was shown on the screen (Figure 5.2) which instructed the subjects to keep their eyes open or closed for 15 s. Totally, 10 trials in eyes open condition and 10 trials in eyes closed condition were recorded randomly and the subjects were informed about the end of each trial by hearing a beep sound.

During calibration runs, a moving right or left hand was shown on the



(a)



(b)

Figure 5.1: Schematic of the experimental design for: (a) experimental group, (b) control group.

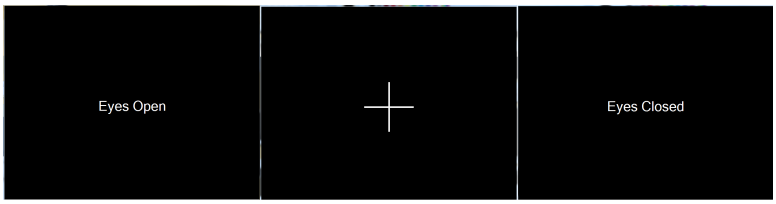
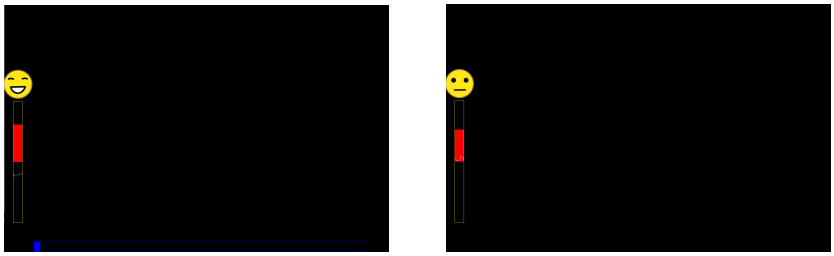


Figure 5.2: Visual cues during resting state recording. A fixation cross is shown on the screen followed by an eyes-opened or an eyes-closed cue.

screen as a visual cue which instructed the subject to perform either right or left hand kinesthetic MI (Figure 5.3). No feedback was provided for the subject during these two runs. Subsequently, two feedback runs were recorded. A real time feedback and a smiley face was shown on the screen to inform the subject whether MI action was correctly detected (Figure 5.4). There was a break period of 2 min between each two consecutive runs.

Subjects performed kinesthetic hand MI in both non-feedback and feedback runs. Each non-feedback run comprised 40 trials of each MI task and lasted about 16 min. Figure 5.5 (a) shows the timing scheme of a non-

Figure 5.3: A visual cue in MI-BCI experiment. The left (right) hand instructs subject to perform left (right) hand MI. The horizontal blue bar showed the preparation time before providing a cue.



(a) (b)

Figure 5.4: The feedback provided for subject. The vertical bar shows real time output of the classifier, smiley face on top of the bar shows: (a) correct MI detection and (b) wrong MI detection.

feedback trial. Each trial lasted about 12 s. It contained a preparatory time segment of 2 s, followed by a visual cue for 4 s, and a 6 s rest period. These two runs were recorded to train a subject-specific MI detection model. The trained model was used later in the next two subsequent runs to detect right versus left hand MI for the same subject.

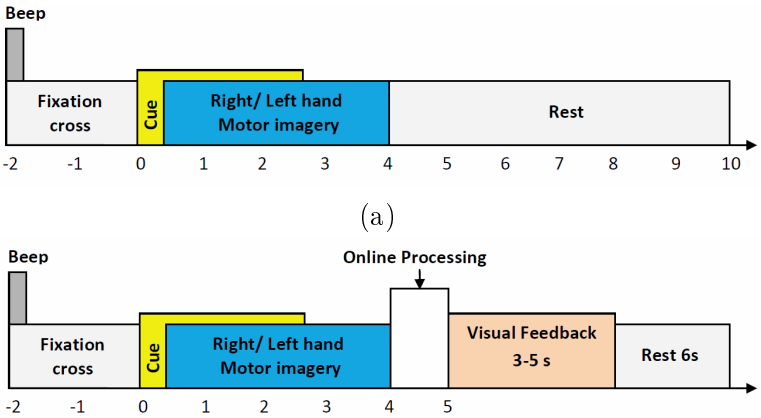


Figure 5.5: Timing scheme of: (a) a non-feedback trial, (b) a feedback trial.

Figure 5.5 (b) shows the timing scheme of an online feedback trial. Each trial lasted about 18 s, a preparatory time segment of 2 s, followed by a visual cue for 4 s. After 1 s of online processing the type of MI task was then detected and a visual feedback was shown for about 3 to 5 s. Each feedback run lasted about 26 min.

5.3.3 Neuro-feedback training (NFT) sessions

The experimental group participated in 3 NFT sessions per week over one month, 12 sessions in total. The NFT sessions aimed to enhance the resting state alpha rhythm (8–13 Hz) of the subjects. Each NFT session was recorded in a separate day for about 20 min. Before and after each training session, 5 min of resting state EEG data were recorded. The instruction for resting state was exactly the same as the the one which is explained in Section 5.3.2. Subsequently, the participants played a game (See Fig. 5.6) for 20 min in which an avatar was moved along a path. The speed of the avatar was controlled by the subjects, and they should keep being relaxed to increase their alpha rhythms and thus speed up the avatar movement. In total, each NFT session took around 30 min and the subjects were instructed to minimize their body limb movement during the whole experiment.

The BCI score shown on the screen was directly proportional to the amplitude of the subject's brain rhythm. The score was between 0 to 100. Relaxation, control of breathing and short-term memory tasks are some of the strategies for enhancing the α-SSD power. Following these strategies increased the BCI score and therefore moved the avatar faster.

Figure 5.6: Experimental setup to collect EEG data during NF training sessions. Subjects should modulate their brain signals to move the avatar on the screen. The BCI score which is shown on the bottom left corner of the screen is the speed of the avatar.

5.4 Methods

5.4.1 SSD-components in resting state

In this experiment, EEG data from 25 channels were recorded using the Nuamps EEG acquisition hardware with unipolar Ag/AgCl electrodes, digitally sampled at 250 Hz with a resolution of 22 bits for voltage ranges of ± 130 mV and band-pass filtered from 0.05 to 40 Hz by the acquisition hardware.

The power of SSD components has been previously proposed as a performance predictor in Section 4.3. The alpha SSD-based components (8–12 Hz) are computed from resting state EEG in both eyes-open and eyes-closed conditions using Eq. 4.1 to 4.7. The neighboring frequency bands are selected as (6–8 and 12–15 Hz). The spatial filter $\mathbf{W}$ derived from Eq. 4.4 is used subsequently for online estimation of SSD components in NFT sessions.

The power of SSD components in eyes open condition is calculated over a 2 s sliding window. The window is shifted 200 ms each time. The 85^{th}

percentile of the resting state powers is then used as upper bound for scaling extracted powers between between 0 to 100. Figure 5.7 shows a sample histogram of resting state α-SSD components power.

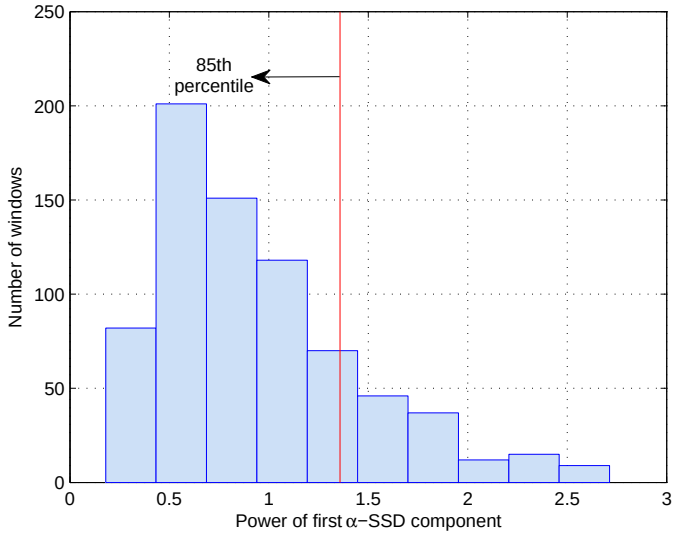


Figure 5.7: Histogram of resting state α-SSD component power in a 2 s sliding window shifted every 200 ms. The $85th$ percentile of the powers are used for scaling the power between 0 to 100.

5.4.2 Online estimation of SSD components in NFT sessions

In order to calculate the α-SSD components in online NFT sessions, every 4 s of the recorded EEG is spatially filtered by the spatial filter $\mathbf{W}$ derived from resting state EEG recorded before each NFT session:

$$\widetilde{\mathbf{X}} = \mathbf{X}\mathbf{W} \tag{5.1}$$

where $\mathbf{X} \in \mathbb{R}^{\tau \times c}$ is the recent recorded EEG signal, $\mathbf{W}$ is the spatial filter derived from resting state EEG data, $\widetilde{\mathbf{X}} = [\widetilde{\mathbf{x}}(1), \widetilde{\mathbf{x}}(2), \cdots, \widetilde{\mathbf{x}}(\tau)]^T$ is the spatially filtered EEG data, where $\widetilde{\mathbf{x}}(t) \in \mathbb{R}^{1 \times n_m}$, n_m is the selected number of SSD filters used for spatially filtering the EEG data ($n_m \leq c$), and τ and

c are the number of samples and channels, respectively.

The spatially filtered $\widetilde{\mathbf{X}}$ is then band-pass filtered. The power $\mathbf{p}_{\widetilde{X}}$ of first two SSD components (n_m=2) are then calculated and averaged over 4 seconds (τ=1000):

$$\mathbf{p}_{\widetilde{X}}(t) = \|\widetilde{\mathbf{x}}(t)\|^2 \tag{5.2}$$

$$\overline{p} = \frac{1}{n_m} \sum_{i=1}^{n_m} \frac{1}{\tau} \sum_{t=1}^{\tau} p_{\widetilde{X}}(t,i) \tag{5.3}$$

Every 200 ms, $\overline{p}$ is computed and then scaled according to mapping scale derived from resting state EEG with eyes open. The scaled value is then smoothed using a moving average over the recent 10 samples and shown as a continuous BCI score on the screen to inform the subject about his performance.

5.5 BCI performance evaluation

In BCI sessions, the collected EEG data from calibration runs were processed using FBCSP algorithm (section 3.5.1) to construct a subject-specific MI detection model. FBCSP band-pass filtered the EEG signal in 9 different bands between 4 Hz to 40 Hz. CSP was applied on the second stage to spatially filter the signal. Two pairs of features from each band were selected. Subsequently, the best 4 discriminative pairs of features were selected using mutual information and then fed into a linear discriminant analysis (LDA) classifier. The trained model was then used in online feedback runs. The output of the trained LDA classifier was continuously computed and shown as a feedback to the subjects. A happy (neutral) face was shown on the screen if the MI action was correctly (wrongly) detected.

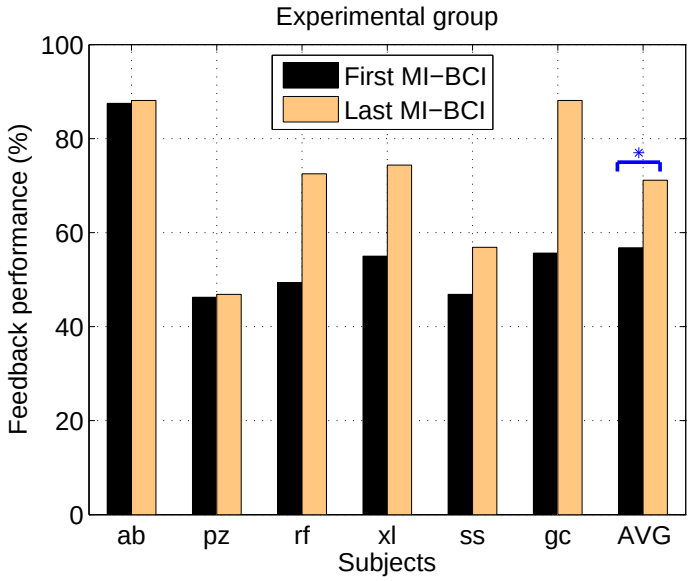


Figure 5.8: The feedback accuracy of the experimental group. (Paired sample t-test: * $p<0.05$)

5.6 Results and discussion

5.6.1 Classification performance of MI-BCI sessions

The feedback performance of the experimental group in two MI-BCI sessions, before and after the NF training sessions, are shown in Figure 5.8. The average feedback performance of the subjects in the first MI-BCI session was 56.77%±15.57 and the performance of the subjects ranged from 46.25% to 87.5%. Only subject '*ab*' who had prior BCI experience achieved high accuracy 87.5% in the first BCI session. The rest of the subjects had accuracy below 70% and thus considered as subjects with BCI deficiency. However, the average accuracy of the experimental group in the second MI-BCI session, which was accomplished after NF training sessions, was considerably increased to 71.14%±16.62.

To find out whether the performances of the subjects in the second session were significantly better than the first session, we conducted a paired sample t-test. The results showed the BCI performance of the subjects were

significantly improved upon their initial accuracies p=0.02. In order to investigate the feasibility of addressing BCI deficiency through NF training, we excluded subject '*ab*', who had high performance in the first BCI session, from our analysis. The results of paired sample t-test revealed the average accuracy of the 5 poor performance subjects significantly (p=0.018) improved from 50.63% in the first BCI session to 67.75% in the second BCI session. Moreover, three participants with BCI deficiency reached accuracy above 70% in their second BCI session. However, the other two participants did not achieved 70% accuracy.

The feedback performance of the control group in two BCI sessions are shown in Figure 5.9. As shown, two out of seven participants had high feedback performance in the first BCI session, while the rest had accuracy below 70%. Thus, both groups had 5 subjects with poor performance. Moreover, the control group, similar to experimental group, performed their second BCI session around one month after their first session. However, the average feedback performance of the control group in the second BCI session showed slightly improvement, which was not statistically significant p=0.41. Only one poor performance subject achieved accuracy above 70%. Moreover, the feedback performances of the two high performance subjects slightly deceased in the second BCI session, but their performances were still around 90%. Although the average performance of the 5 subjects with BCI deficiency from the control group was increased from 56.75% to 62.05%, the improvement was not statistically significant p=0.14.

5.6.2 Neurofeedback results

The resting state alpha power of subjects with BCI deficiency from the experimental group over the 12 NF training sessions are shown in Figure 5.10. Each dot represents the average relative resting state alpha activity in each

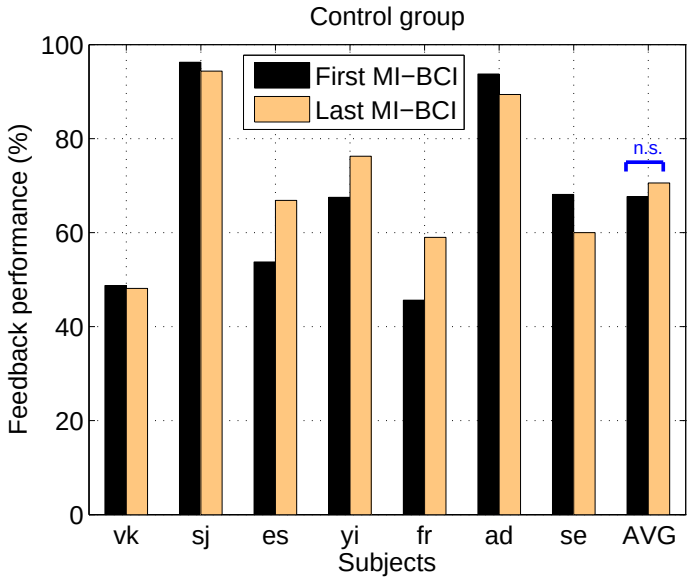


Figure 5.9: The feedback accuracy of the control group. (Paired sample t-test: n.s. p>0.05)

NF session. The relative alpha power was calculated by dividing the alpha band power in the range 8–13 Hz by the broad-band power in the range 4–30 Hz. As can be seen, the resting state alpha activity increased consistently over time. The resting state alpha of poor performance subjects in the first and last training session were compared in Figure 5.11(a). Conducting a paired sample t-test revealed that the alpha power in the $12th$ training session was significantly higher than the first training session p=0.04. Figure 5.11(b) also revealed that for all subjects the alpha power in the $12th$ session was higher than the first training session.

In order to better analyze the time course of the resting state alpha power over the NF sessions, we performed regression analyses for the subjects with BCI deficiency. The predictor variable was the NF session number and the dependent variable was the resting state alpha power. The derived regression model showed significant trend ($F(1,10)$=21.65, p=0.001) and explained 68% of variance of the resting state alpha power over the NF training sessions

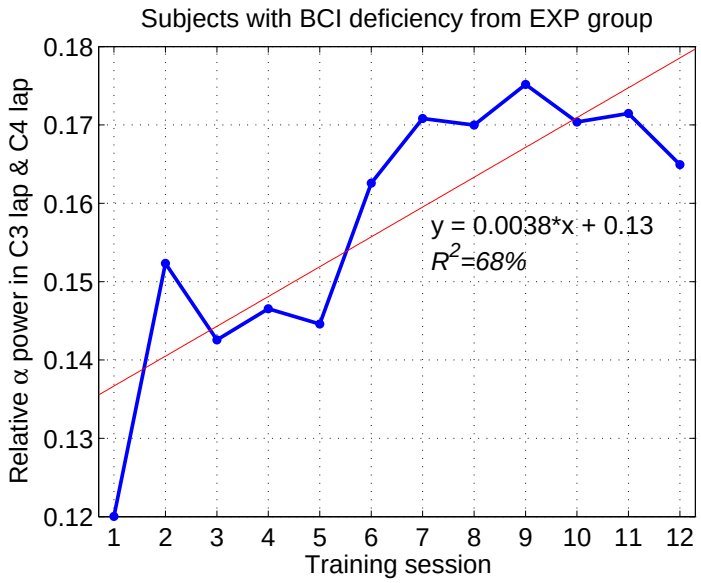


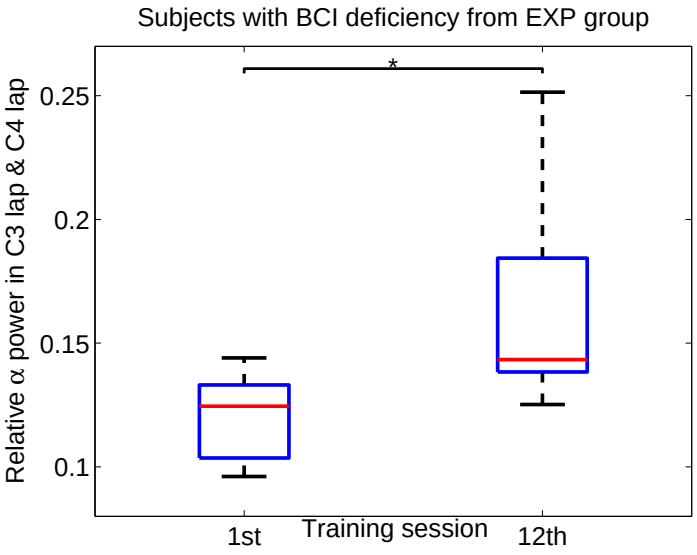
Figure 5.10: Average relative resting state alpha power of subjects with BCI deficiency over the 12 NF training sessions.

(R^2=0.68).

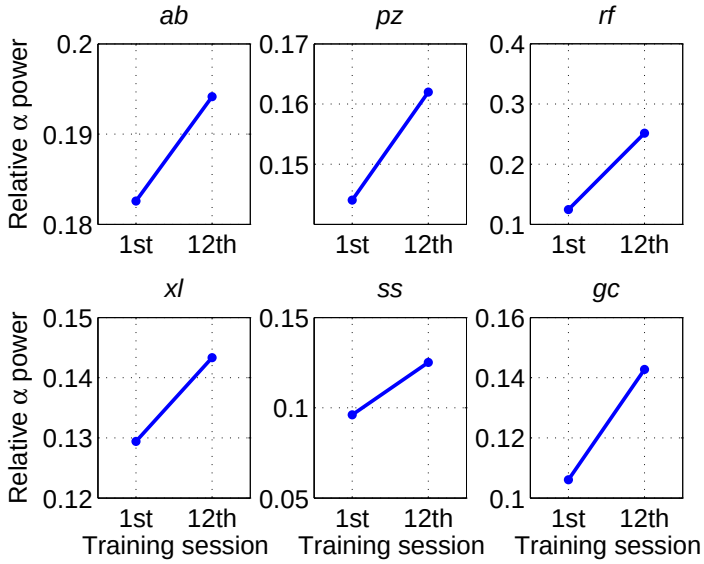
In contrast, the regression model for the high performance subject did not reveal significant trend ($F(1,10)$=1.52, p=0.247) and this subject showed no prominent changes in the resting state alpha power.

5.7 Summary

In this chapter, we proposed an approach to address BCI deficiency, which is one of the challenges in BCI applications. We hypothesized that by enhancing the resting state alpha rhythm, the MI-BCI performance can be improved. Therefore, we conducted an experiment to train the subjects increase their alpha rhythm in 12 NFT sessions. Relaxation was a successful mental strategy to enhance the resting state alpha in NFT sessions. The results showed that NFT helped subjects significantly improve their resting state alpha rhythms over the sensorimotor area. Moreover, the MI-BCI performance of the subjects significantly improved after NFT sessions. Therefore, it can



(a)



(b)

Figure 5.11: Comparing the relative alpha power of subjects in the first and last NF training session. (a) Boxplot of the relative alpha power for participants with BCI deficiency. (Paired sample t-test *p<0.05); (b) The relative alpha power of each subject in the first and last training sessions.

be inferred that the BCI performance of the subjects improved by enhancing their resting state alpha rhythm. In conclusion, NF can be considered as a promising method to alleviate BCI deficiency of the subjects.

Chapter 6

KL Distance Weighting of FBCSP to reduce Inter-Session Non-stationarity

6.1 Introduction

EEG-based BCI is operated by voluntary modulation of SMRs in the alpha, beta, or both frequency bands. Motor imagery (MI) or imagination of movement can modulate the SMRs [24] and thus MI is one of the commonly used methods for controlling an EEG-based BCI system.

Several MI tasks such as imagination of hand, foot, and swallow can be used for modulating the SMRs. However, the performance of the subjects in each of these imagery tasks are different. Therefore, in some experiments [33, 35] the subjects were instructed to perform different types of imagery tasks during calibration session. Those tasks with higher performance are then chosen for training the BCI.

Hand imagination in comparison with other imagery tasks are more used in BCI. The brain patterns during right and left hand MI are different [78]. During imagination of right hand movement, the cortical networks under the sensorimotor cortex in the left hemisphere desynchronizes, while during left hand motor imagery desynchronization is in the right hemisphere. Typically,

the desynchronization is more significant around C3 and C4 channels. These changes in cortical activity are detected by proper EEG processing which was previously explained in Section 2.3.3.

Typically, a user imagines his own hands movement while performing hand MI. [93, 109, 165]. Although various strategies such as finger tapping or wrist twisting are provided for the subjects to perform MI, several subjects are still unable to perform MI correctly. Consequently, they cannot modulate their SMRs. On the other hand, some subjects may change their MI strategy during the experiment, since they may have an impression of doing MI wrongly. Therefore, these subjects may not perform MI task consistently during a single session of BCI. These problems lead us to think of a better solution for calibrating an EEG-based MI-BCI.

Motor Execution (ME) is one of the possible activities which can be used for calibration of a BCI system [166]. However, the final goal of BCI is to help people who may have severe motor disabilities. These disabled people may not be able to perform actual movement. Therefore, it would be better to have a more general solution for calibrating MI-BCI. Passive Movement (PM) is another possible solution which has been applied successfully for calibrating MI-BCI for both group of healthy [47] and disabled subjects [15]. More details about calibration of MI-BCI with PM are provided in Section 6.2.

Apart from the advantages of calibration of MI-BCI with PM, there might be a difference between feature distribution of PM and MI. The difference between feature distribution of calibration session and evaluation session is not only because of calibrating with PM. Generally, the inherent non-stationarity characteristics of EEG signal result in difference between feature distribution of calibration and evaluation session. Therefore, we need to apply an algorithm which can address inter-session non-stationarity.

In this chapter a new adaptive algorithm to address inter-session non-stationarity is proposed. The proposed method iteratively updates the feature space in order to reduce the inter-session difference of calibration session and evaluation session. The proposed method is explained in Section 6.3.

6.2 EEG-based MI-BCI calibration with passive movement (PM)

PM refers to moving a part of the body by an external force such as a robot, haptic knob, or therapist rather than its own movement [96]. PM does not need any voluntary or imaginary motor action. Therefore, both healthy and disabled subjects can take advantage of PM. Moreover, data collection during PM is more facilitated since subjects may rarely get fatigue, and consequently more trials of the EEG can be recorded.

Studies have shown that MI, ME and PM activate the motor system similarly [9, 167, 168]. Therefore, the ERD/ERS patterns, which are generated by each of these activities, are quite similar in both mu and beta frequency bands over motor area. This finding proves that the MI-BCI can be calibrated with either PM or ME data. It has been shown that foot ME can be used for training a MI-BCI system [87, 166] for MI detection. The ERD/ERS patterns of foot MI and foot ME were similar. In other words, not only the ERD patterns during the movement were similar but also the beta rebound or ERS patterns generated after the movement were quite identical. However, as mentioned earlier in 6.1, ME is not practical for MI-BCI calibration in most clinical applications.

The feasibility of calibrating the MI-BCI with PM was also studied in [169]. It has been shown that calibration with PM resulted in higher model accuracy and off-line session-to-session transfer than calibration with PM.

The main advantage of setting up the initial classifier with PM is shortening the setup time needed for online BCI. A naive subject does not need to go through any training which could be intensive. Therefore, PM can be used for calibration of MI-BCI in practical applications. As an example, PM calibration facilitates BCI-based stroke rehabilitation.

6.2.1 PM in BCI-based stroke rehabilitation

MI-BCI has been shown to be used for neurorehabilitation [15, 17]. Subjects with different level of disabilities can perform MI, since it does not need any actual movement. Moreover, MI can access the motor system at all stages of stroke recovery [16]. However, some subjects cannot perform MI correctly and consistently.

It was shown previously that MI-BCI was calibrated with PM for stroke patients therapy [9, 15]. PM generates more consistent data for calibrating the BCI. Moreover, the PM data collection is easier than MI, because the subjects do not engage in any mental task, and thus they don't get tired easily. PM can be conducted by a mechatronic finger rehabilitation device [9], or a MIT-Manus robot [15]. Both studies showed the feasibility of calibrating MI-BCI with PM for stroke patients. Kaiser et al. in [9] also calibrated the MI-BCI with ME. However, performing ME is not possible for all the subjects since stroke patients have different level of disability and motor impairment.

6.2.2 Non-stationarity in session-to-session transfer

The non-stationarity may cause because of: 1) inter subject variability and changes in the subject's brain processes within or across sessions such as changes in task involvement or fatigue, 2) physiological artifacts such as body movement, swallowing or blinking, 3) instrumental artifacts such as changes in impedance of electrodes or changes in their placement across sessions

The statistical characteristics of EEG signal from calibration session to evaluation session may be changed and therefore results in a difference between the feature distribution of the calibration and the evaluation session. The non-stationarity may cause because of: 1) inter subject variability and changes in the subject's brain processes within or across sessions such as changes in task involvement or fatigue, 2) physiological artifacts such as body movement, swallowing or blinking, 3) instrumental artifacts such as changes in impedance of electrodes or changes in their placement across sessions.

The inherent non-stationarity behavior of the EEG signal can deteriorate the performance of the session-to-session transfer. Several research have been done to address inter-session non-stationarity. General speaking, adaptation is the most commonly used method to address non-stationarity. One category of the proposed methods is based on adapting the classier. This group of methods are briefly reviewed in Chapter 7.

The second category of the proposed adaptive methods to address inter-session non-stationarity is based on adapting the feature space. These methods are based on updating the train model using adaptation. Covariate Shift adaptation is a successful method to overcome the difference between train and test data and thus it is a suitable method to be used for session-to-session transfer [170, 171]. Domain space adaptation [106] is another proposed method for session-to-session transfer. This is a preprocessing method applied before CSP. Some other studies proposed methods for adapting the CSP and thus adaptively updating the features [44, 172]. By extracting the updated features, the classifier can be retrained.

Joint feature extraction and classification has been also proposed in some previous studies [43, 113]. Expectation Maximization (EM) algorithm was used iteratively to extract the features and retrain the classifier. Their

method was based on semi-supervised learning since the train data was augmented with the predicted labels of the current evaluation data.

All above mentioned studies address inter session non-stationarities. However, the non-stationarities may also appear within a session as well. Several methods has also been proposed for addressing within session non-stationarity which are not the focus of this chapter. For example, KLCSP [40] and sCSP [173], which find invariant features, or stationary subspace analysis which finds the stationary parts of the EEG signal [39] are some of the examples proposed for tackling within session non-stationarities.

Calibrating MI-BCI with PM may cause the feature distribution of calibration session and evaluation session to be different. Therefore, the performance of the subjects in session-to-session transfer may decrease. Figure 6.1 shows two dimensional feature space for a subjects with high performance in calibration session and low performance in session-to-session transfer.

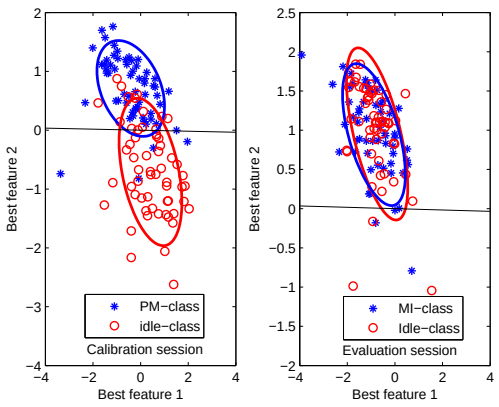


Figure 6.1: Two dimensional feature space of a subject *kk* with high performance accuracy in calibration session (PM) but low performance accuracy in evaluation session (MI). MI/ PM class is shown by (*) and idle class is shown by (o)

As can be seen in the figure, the left plot shows the feature space of calibration session (PM). The right plot shows the feature space of evaluation session (MI). As shown, there is a shift in feature space. The two classes

are not linearly separable and the initial trained hyper plane for calibration session is not optimal for evaluation session. Adaptive methods may be useful for addressing such non-stationarities between PM and MI sessions.

6.3 Iteratively Updating FBCSP Using KL Distance Weighting

To address inter session non-stationarity, a new adaptive algorithm is proposed in this chapter. In our proposed algorithm, the features are extracted using FBCSP [118] and iteratively updated based on a newly recorded batch of EEG data. The proposed method can be implemented in on-line scenario. In the following, first the CSP [104] and FBCSP [118] are briefly reviewed, then the proposed method is explained in Section 6.3.3 and 6.3.4.

6.3.1 Common Spatial Pattern (CSP)

EEG preprocessing is a crucial part of MI-BCI system. Among different proposed methods for preprocessing, CSP algorithm [88] serves as an effective tool for discriminating the two classes of EEG data by maximizing the variance of one class while minimizing the variance of the other class.

The CSP algorithm computes the spatial filter (i.e., the transformation matrix $\mathbf{W}$) by solving the eigenvalue decomposition problem. The main computational steps of CSP are explained in the following:

$$\mathbf{Z}_i = \mathbf{W}^T \mathbf{E}_i, \tag{6.1}$$

where $\mathbf{E}_i \in \mathbb{R}^{c \times \tau}$ denotes the i^{th} band-pass filtered EEG trial ; $\mathbf{Z}_i \in \mathbb{R}^{c \times \tau}$ denotes $\mathbf{E}_i$ after spatial filtering, $\mathbf{W} \in \mathbb{R}^{c \times c}$ denotes the CSP projection matrix; c is the number of channels; τ is the number of EEG samples per channel; and ^{T} denotes the transpose operator.

The CSP algorithm computes the transformation matrix $\mathbf{W}$ by solving the eigenvalue decomposition problem:

$$\mathbf{\Sigma}_1 \mathbf{W} = (\mathbf{\Sigma}_1 + \mathbf{\Sigma}_2)\, \mathbf{W} \mathbf{D}, \tag{6.2}$$

where $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ are estimates of the covariance matrices of the band-pass filtered EEG of the respective motor imagery action, $\mathbf{D}$ is the diagonal matrix that contains the eigenvalues of $\mathbf{\Sigma}_1$.

The m first and last rows of $\mathbf{Z}_i$ i.e. $\mathbf{Z}_{i_p}$, $p \in \{1 \cdots 2m\}$ forms the feature vector $\mathbf{v}_i$ given in Eq. 6.3 as inputs to a classifier.

$$\mathbf{v}_i = log(\frac{var(\mathbf{Z}_{i_p})}{\sum_{p=1}^{2m}(\mathbf{Z}_{i_p})}) \tag{6.3}$$

6.3.2 Filter Bank Common Spatial Pattern (FBCSP)

It has been shown the performance of the spatial filters constructed by CSP algorithm are dependant on their operational frequency band [88]. In other words, choosing the subject specific frequency band before applying CSP enhances the performance of the users. This is mainly because the ERD/ERS of SMRs for each subject has a specific spectral pattern. Therefore, applying a method for selecting the subject specific frequency band is desirable. FBCSP algorithm [117, 118, 174] is one of the proposed methods which automatically selects the key temporal spatial discriminative EEG characteristics. Hence, due to the subject-specific frequency band selection by FBCSP higher performance is achieved in comparison with normal CSP [117]. In this chapter, we extracted features using FBCSP. Figure 6.2 shows the structure of the FBCSP.

As explained earlier in Section 3.5.1, FBCSP has four progressive stages. On the first stage it band pass filters the EEG signal in 9 different bands,

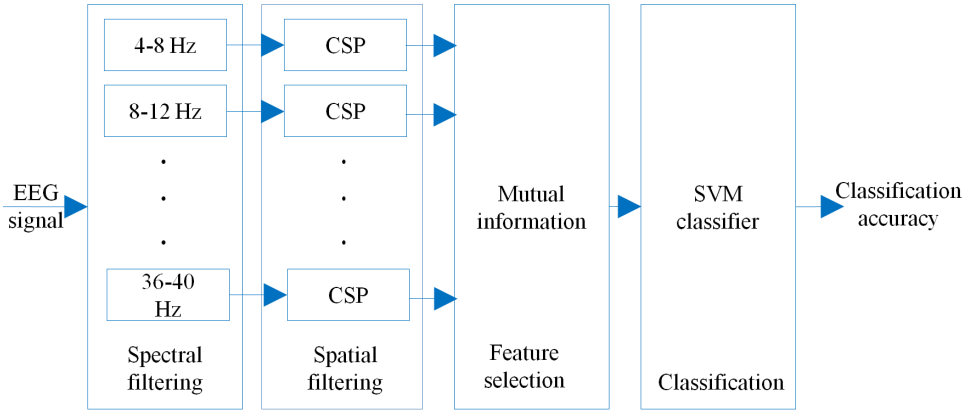


Figure 6.2: Architecture of filter bank common spatial pattern (FBCSP) algorithm.

namely, 4-8, 8-12, ... , 36-40 Hz. Then second stage employs the CSP to spatially filter the signal. For each band m pairs of features are selected. Therefore, the total number of features is $2m \times 9$. On the third stage, the best k pairs of discriminative CSP features are selected using MIFBIF algorithm. The value of k is selected using 10x10 cross-validations on the calibration data. Some previous studies used k=4 [117, 118]; however, since the CSP features are paired, the corresponding pair of features is also included if it is not selected. Therefore, the final number of selected features will be between 4 and 8. On the last stage, the selected features are fed into SVM classifier. The final accuracy shows how well the subjects can discriminate MI versus idle state. More details about the FBCSP algorithm can be found in [117, 118, 174].

The best number of CSP features is selected using 10x10 cross-validations on the calibration data. Moreover, some previous studies , used k=4; however, since the CSP features are paired, the corresponding pair of features is also included if it is not selected. Therefore, the final number of selected features will be between 4 (if all 4 features selected are from 2 pairs of CSP features) and 8 (if all 4 features selected are from 4 pairs of CSP features)

6.3.3 KL Distance Weighting of FBCSP

Although selecting the subject specific frequency band enhances the performance of the users, adaptation is still required to overcome the inter session non-stationarity. In this chapter, we proposed a new adaptive method for updating the extracted features. The CSP of each band is updated iteratively. A limited number of EEG trials from the evaluation session is used for updating the features. The update of CSP is based on update of the covariance of each class in each band:

$$\mathbf{\Sigma}_{b,(\omega)} = (1-\alpha)\mathbf{\Sigma}^{tr}_{b,(\omega)} + \alpha\mathbf{\Sigma}^{ts}_{b,(\omega)}, \tag{6.4}$$

where $\mathbf{\Sigma}^{tr}_{b,(\omega)}$ denotes the estimate of the covariance matrix of class $\omega \in \{1,2\}$ in b^{th} band, and tr and ts denote the covariance of train and evaluation data, respectively.

The weighting parameter α is defined as $\alpha = \frac{n}{N+n}$, where N is the total number of trials of passive movement calibration data and n is the number of recorded trials from evaluation session which we aimed to use for updating the covariance. This idea is similar to updating CSP in Composite Common Spatial Pattern (CCSP) [175]. In CCSP the CSP is updated by using the data from other subjects. Here, the CSP on each band is updated using the newly recorded data from the evaluation session.

In order to better estimate the covariance matrix of each class, we proposed a new method for updating the covariance matric. Eq. 6.5 is a new covariance estimation method that includes a new KL-distance [176] weighting term. In fact the new estimation has a similar structure to Eq. 6.4 but assigns some extra weights to the newly recorded trials from the evaluation data:

$$\mathbf{\Sigma}_{b,(\omega)} = (1-\alpha)\mathbf{\Sigma}^{tr}_{b,(\omega)} + \frac{1}{KL_{b,(\omega)}}\alpha\mathbf{\Sigma}^{ts}_{b,(\omega)}, \tag{6.5}$$

where $KL_{b,(\omega)}$ is a KL distance that shows the difference between probability distribution of the train and incoming trials from evaluation data, and it is defined as follows:

$$KL_{b,(\omega)} = 0.5\{\log(\frac{\det \mathbf{\Sigma}^{tr}_{b,(\omega)}}{\det \mathbf{\Sigma}^{ts}_{b,(\omega)}}) + \quad trace((\mathbf{\Sigma}^{tr}_{b,(\omega)})^{-1}\mathbf{\Sigma}^{ts}_{b,(\omega)}) - D\}, \qquad (6.6)$$

where D is the dimension of the covariance matrix and *det* represents determinant of a matrix.

6.3.4 Batch mode updating of KL weighted FBCSP

In order to update the FBCSP with KL distance weighting, we need to use recently recorded data from current evaluation session. A batch mode semi-supervised learning method is proposed to update the FBCSP. Semi-supervised learning is one type of adaptive learning methods dealing with these cases where both labeled and unlabeled data are available [177].

In online MI-BCI system, the subject may be instructed to perform specific MI action; thus the labels of the evaluation data are known. However, in real BCI applications it is more preferable that the subject be free to perform either types of MI task. In such later case the labels are unknown, since the subject is not instructed to perform any specific MI task. Here in this thesis, we only focused on synchronous BCI where the labels of the tasks are known.

Figure 6.3 illustrates the schematic diagram of our proposed online batch mode semi-supervised algorithm. As shown, the algorithm needs some settings from offline calibration session. Labeled PM data from offline calibration session are used to train the classifier. During online session, MI is detected based on the trained classifier which is updated iteratively. A new batch of recorded EEG data from online session is used to update FBCSP.

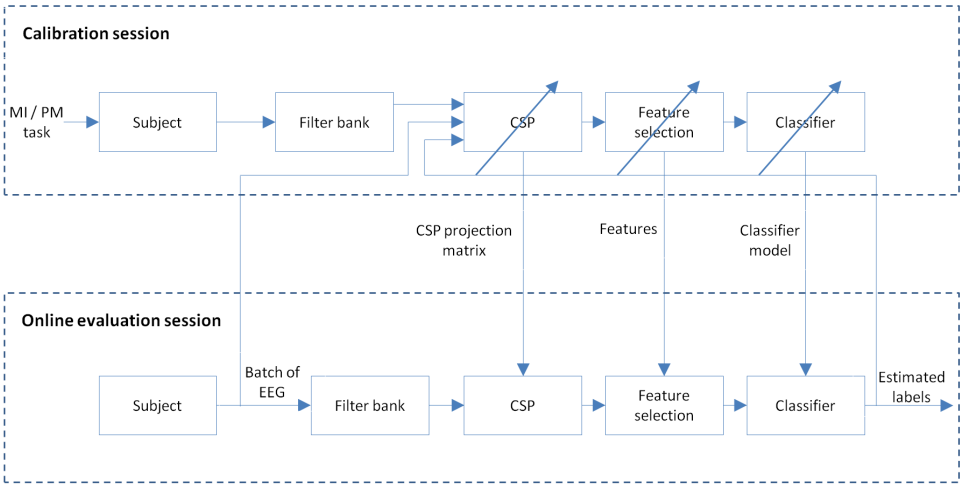


Figure 6.3: Architecture of online batch mode semi-supervised learning method based on FBCSP algorithm.

Algorithm 1 describes how the proposed method works [47]. The maximum number of iterations K and also the number of EEG trials used for updating the algorithm are selected through several simulations, that can be also selected according to some previous literature [43].

Algorithm 1 Batch mode updating FBCSP with KL distance weighting

1: Train FBCSP using the calibration data (offline).
2: Choose the number of trials in each batch.
3: Choose the maximum number of iterations K for updating the labels.
4: Estimate the labels of the recent batch of data from evaluation session with current trained model.
5: Keep the estimated labels for calculating the accuracy performance.
6: **for** k=1:K: **do**
7: Add the newly recorded batch of evaluation data with estimated labels to the train data.
8: Retrain FBCSP to derive the new features, and re-estimate the labels of the newly recorded evaluation data.
9: **end for**
10: Update the train model by adding the newly recorded evaluation data with the estimated labels after K iteration.
11: Go to step 4

6.4 Experimental set-up

Here, the EEG data recorded from the 15 healthy subjects are used. Two of the subjects were left handed and therefore performed MI and PM by their left hand while the rest performed by their right hand. All the subjects were asked for ethics approval and informed consent. EEG signal were recorded by using the Nuamps EEG acquisition hardware (http://www.neuroscan.com) with unipolar Ag/AgCl electrodes channels, digitally sampled at 250 Hz with a resolution of 22 bits for voltage ranges of 130 mV. EEG recordings from all 27 channels were bandpass filtered from 0.05 to 40 Hz by the acquisition hardware. All subjects were instructed to minimize their physical movement and eye blinking throughout the EEG recording process.

The two sessions of EEG data from each subject were collected on two separate days. During the first session, four non-feedback runs were recorded. The first two runs collected EEG from a subject while performing MI of the chosen hand versus idle state. During these two runs, the subjects were instructed to perform kinaesthetic MI of their chosen hand right after a visual cues displayed on the computer screen in each trial. During the idle state, the subjects were instructed to perform mental counting. This instruction was given to define the idle state to the subject. During the next two runs EEG data was collected from the subject while performing PM of the chosen hand with the haptic knob robot [123] and idle state. During these two runs, the subjects were supposed to be relaxed while the movement of the chosen hand was performed using the haptic knob robot [123]. During the idle state subjects performed mental counting similar to the first two sessions.

Each run lasted for about approximately 16 minutes that comprised of 40 trials of either MI or PM, and 40 trials of idle state condition. Each trial comprised a preparatory segment of 2s, the presentation of the visual cue

for 4s, and a rest segment of at least 6s. Each trial lasted approximately 10s, and a break period of at least 2 minutes was given after each run of EEG recording. To calibrate the subject-specific model from performing MI (PM), the first (last) two runs were used.

On the second session, four runs of EEG data were collected with feedback from the subjects while performing MI of the chosen hand versus idle state. In the first two runs of the second session the BCI system was calibrated by MI calibration session data, and in the next two runs the BCI was calibrated by PM calibration session. Each run again lasted for about approximately 16 minutes that comprised of 40 trials of motor imagery and 40 trials of idle state.

6.5 Results

6.5.1 Classification results of MI and PM calibration session

Figure 6.4 shows the baseline classification accuracy results of the subjects using FBCSP and SVM classifier. The 10×10-fold CV accuracy of both MI and PM calibration sessions were compared. As shown, the averaged performance of subjects was 73.59%±16.29% in MI calibration session and 75.78%±16.36% in PM calibration session. More specifically, the results of three subjects *at*, *hh* and *ad* showed that the averaged accuracies of detecting MI versus idle state (59.63%) was much lower than the averaged accuracies of detecting PM versus idle state (70.19%). A paired t-test was conducted to compare accuracy of subjects in MI and PM sessions. Although the average performance of subjects in PM session was higher than MI session, there was no significant difference in accuracy of subjects $t(14)$=-1.44, p=0.17.

Six of the subjects (i.e., *hh*, *kk*, *ks*, *s*, *zy*, *lj*) had some prior experience in operating MI-BCI. The remaining nine were BCI-naive subjects. Based on

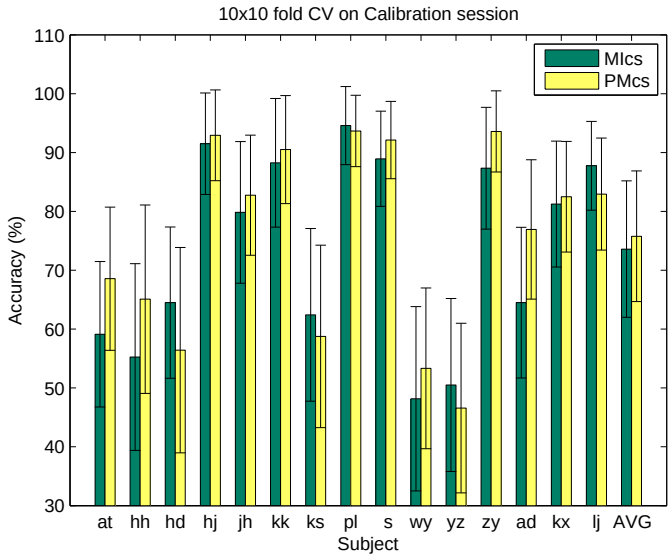


Figure 6.4: 10×10-fold cross-validation accuracies of MI and PM calibration sessions (denoted MIcs and PMcs). Errorbars indicate the standard deviation over cross-validation folds.

this information, the results shows that the averaged accuracies of detecting MI (PM) from idle state for naive subjects were lower 70.44% (78.32%) than the experienced subjects 72.63% (80.50%). Nevertheless, four BCI-naive subjects *jh*, *hj*, *pl* and *kx* had accuracies greater than 80% in both MI and PM sessions and only one experienced subject had accuracy less than 70% in both MI and PM session.

6.5.2 Visualization of the spatial filters for MI and PM calibration session

Figure 6.5 and Figure 6.6 compare the CSP patterns of MI and PM for selected subjects with high and low performance. Figure 6.5 shows the first CSP pattern and filter computed from MI and PM calibration sessions for four subjects *kk*, *pl*, *s*, and *zy* who had mean CV accuracy above 80%. For each subject the first CSP pattern and filter corresponding to the largest eigenvalue is plotted. The first row of results shows the MI pattern and

filter, and the second row shows the results of PM. Note that the colormap has no association with the sign of the patterns and filters. As can be seen in this figure, the MI and PM patterns are very similar.

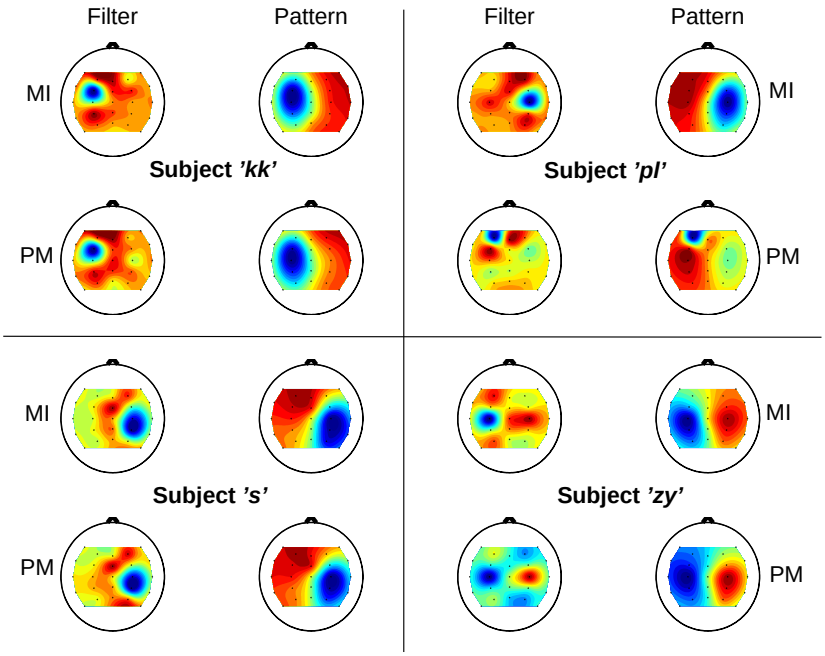


Figure 6.5: The first CSP pattern and filter for both MI and PM task for subjects *kk*, *pl*, *s*, *zy* who had high performance in calibration session. The color-scale is in arbitrary units.

Figure 6.6 shows the patterns and filters of subjects *hh*, *ks*, *wy* and *ad* with low performance in calibration session. By comparing Figure 6.5 and 6.6, it can be seen that there is less similarity between MI and PM for low performance subjects. No hand MI pattern was detected over motor area for subject *ad*, while during PM there was a desynchronization around C4 channel. Nevertheless, the MI and PM patterns for subject *ks* were somehow similar.

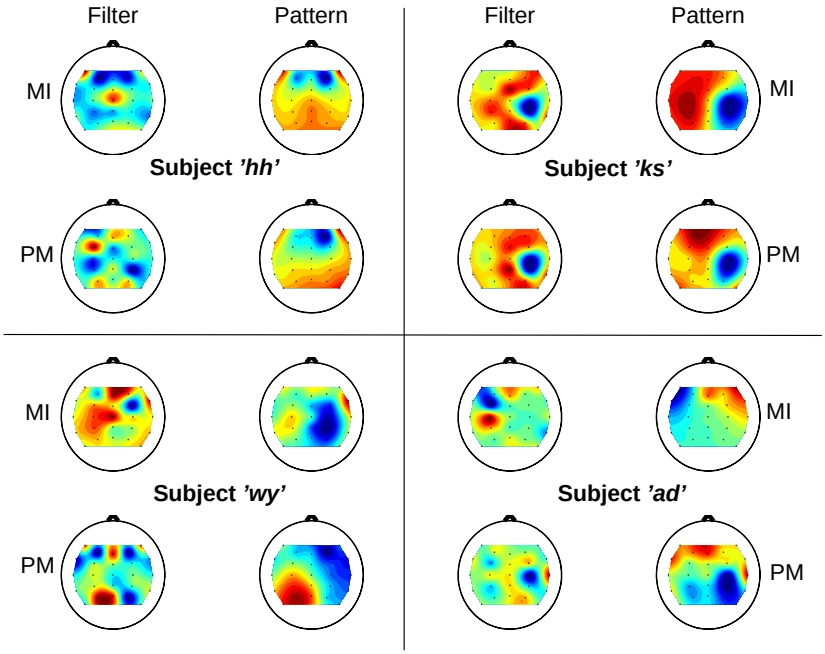
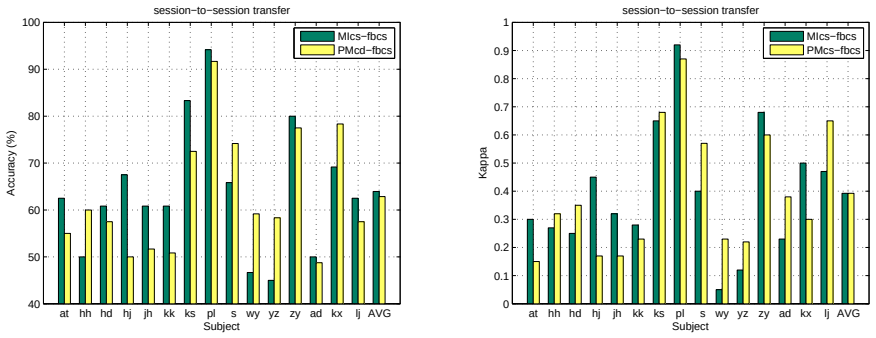


Figure 6.6: The first CSP pattern and filter for both MI and PM task for subjects *hh*, *ks*, *wy*, *ad* who had low performance over MI and PM calibration session. The color-scale is in arbitrary units.

6.5.3 Session-to-session calibration session

In this section the performance of the subjects in session-to-session transfer is evaluated. The results are based on both MI and PM calibration sessions. Figure 6.7 shows the baseline results in session-to-session transfer when FBCSP and SVM classifier are used for evaluating the performance of the subjects in online evaluation session. Figure 6.7(a) compares the accuracies of the online session when calibrated by MI and PM. As shown, the averaged performance are 63.94% and 62.86 %, respectively. The performance of the subjects were also estimated in terms of maximum Kappa value [178]. The Kappa value was computed from two seconds before the cue time to 2 seconds after cue time for every point in time across all the trials of the evaluation session. The results shows calibrating MI-BCI with

MI and PM yielded similar maximum Kappa value (0.39).



(a) (b)

Figure 6.7: Offline session-to-session transfer performance of the subjects in detecting MI versus idle state when MI-BCI is calibrated by MI or PM (denoted MIcs-fbcs and PMcs-fbcs) using FBCSP and SVM classifier. The performance of the subjects in online feedback session are shown in terms of (a) accuracies and (b) maximum Kappa values.

On the next step the proposed method, known as online batch mode semi-supervised learning with KL distance weighting (KLBM-FBCSP), was applied to evaluate the performance of the subjects in evaluation session. According to our simulations and also similar to some other literature [43], the number of iterations were fixed to K=3. Due to the total number of test trials we planned to record and the speed of the online system, we choose 10 trials in each batch. The effectiveness of our proposed method was compared with baseline FBCSP. Figure 6.8 compares the results of these two methods. The average accuracy over 15 subjects using KLBM-FBCSP (64.15%) which was higher than average accuracy of FBCSP (62.86%).

The results of using KLBM-FBCSP also indicated that 10 out of 15 subjects achieved higher accuracy in comparing to baseline FBCSP. Analyzing the results of KLBM-FBCSP showed that the averaged accuracies of experienced subjects (66.54%) was higher than the accuracies of BCI-naive subjects (62.55%). Nevertheless, performing an independent-sample t-test showed no significance difference between these two groups $t(13)$=-0.52, p=0.61. The

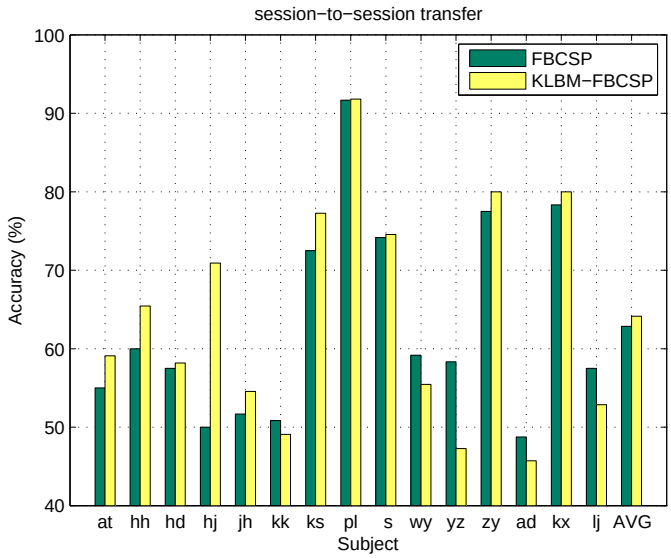


Figure 6.8: The accuracies of subjects in MI-BCI evaluation session when calibrated by PM data using FBCSP and online batch mode semi-supervised method with KL distance weighting (KLBM-FBCSP).

accuracies of the subjects using KLBM-FBCSP when calibrated by PM and MI were also compared. Figure 6.9 shows the comparison results. As can be seen, applying our proposed method for online MI detection resulted in higher performance when calibration was performed by PM (64.15%) than MI (62.61%).

6.6 Discussion

In this work, the feasibility of calibrating MI-BCI with PM in online scenario was explored. It has been previously shown that PM can be used for calibrating MI-BCI [169]. Although no significant improvement was achieved, the averaged performance of MI-BCI calibrated by PM was higher. However, to the best of our knowledge no study has been proposed to take the advantage of using PM data for adapting the BCI system.

The CV accuracies of MI and PM calibration session showed higher av-

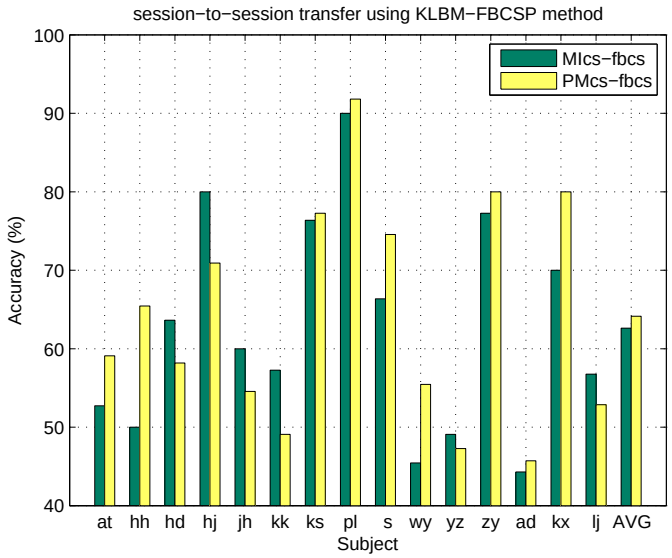


Figure 6.9: The accuracies of motor imagery (with feedback) detection using passive movement (denoted PMcs-fbcs) or motor imagery (denoted MIcs-fbcs) for calibration using online batch mode semi-supervised method with KL weighting.

eraged accuracy for PM than MI which was in line with the findings of previous studies [9, 169]. Although some subjects such as *hj* performed MI quite well even in their first experience with MI-BCI, it is generally expected that BCI-experienced subjects have better MI-BCI performance. 10×10-fold CV accuracies of PM showed that the averaged accuracy of BCI-experienced subjects was around 8% higher than the averaged accuracies of BCI-naive subjects.

Studies have shown performing MI and PM results in similar brain patterns [9, 167, 168]. This finding leads researchers to the idea of calibrating the MI-BCI with PM. Comparing the CSP patterns and corresponding filters of MI and PM in Figure 6.5 and 6.6 showed that the similarity of the patterns for high performance subjects were more than low performance subjects. It may be concluded that the selected low performance subjects except *ks* were not able to modulate their brain rhythms. Subject *ks* was the only subject with low calibration session accuracy (<60%) but higher evaluation session

accuracy (>70%). It may be inferred that this subject was able to modulate his brain rhythms but the rhythms were not detected by the algorithms. One possible reason can be because the subject did not perform motor imagery well and consistence, so the rhythms' modulations in the calibration session were not detected as good as in the evaluation session. It might also be happened because the subject's state or mood was not ideal in the calibration session. Therefore, the performance of the subject in the calibration session was lower than the evaluation session.

The offline analysis of session-to-session transfer (Figure 6.7) showed that on average the detection of MI versus idle state when the model is trained by MI and PM were similar. To have a better comparison maximum Kappa values were also compared. Kappa coefficient is a continuous measure computed over the longer period of time from -2 to 2 second. Evaluating the performance of the subjects using maximum Kappa values also resulted in a similar performance for both conditions. It may be concluded that the MI and PM sessions are quite comparable and resulted in a similar session-to-session transfer performance. Consequently, PM can successfully replace MI during the calibration session. As stated earlier, PM facilitates data collection during calibration session since less effort is required by subject and thus the subject may get fatigued later during one long session of BCI.

Comparing the proposed adaptive method (KLBM-FBCSP) with non-adaptive baseline method (FBCSP) showed that adaptation in our problem was helpful to alleviate the difference between the calibration session and the evaluation session. Using KLBM-FBCSP calibrated by PM improved online MI detection for 10 out 15 subjects. As an example, subject *hj* who had a chance level performance in offline analysis, shows about 20% improvement by using online adaptation. For the rest of the subjects, the drops could be still because of the difference between the MI and PM calibration sessions.

6.7 Summary

There is a common issue on most MI-BCI systems. It is not possible for all the subjects either disabled or healthy to perform MI correctly. One of the feasible proposed methods to overcome this challenge is to calibrate the system with PM data. However, inter-session non-stationarity may still exist and adversely affect the performance of the subjects. In this chapter we proposed a new semi-supervised method for adapting the feature space and thus overcome the possible inter-session non-stationarity. Therefore, a MI-BCI system calibrated by PM was used for online MI detection with feedback.

The results showed that on average the accuracies of the MI-BCI system calibrated by PM data in online system is slightly better than the one calibrated by MI in both offline and online systems. The results may improve more by using some advanced adaptation methods to overcome the difference between calibration sessions of MI and PM. However, the results were promising enough to suggest applying the proposed method for online MI detection in medical applications.

Chapter 7

Adaptive Extreme Learning Machine to Enhance Session-to-Session Transfer

7.1 Introduction

Some of the unsolved challenges in EEG-based BCIs may adversely affect the BCI users' performance. One of the most challenging problems is non-stationarity of EEG signal which may occur due to several factors such as [179]: 1) Intra subject variability usually happens because of the changes in subjects state and mood over different sessions or even within a session; 2) Physiological artifacts are generated by the user himself through eye blinking, muscle movement or respiratory; 3) Instrumental artifacts are originated due to the changes in electrodes position or their impedance during EEG recording.

Typically in MI-BCI experiment, a calibration session is required to train a model for MI detection in an evaluation session. Statistical characteristics of EEG vary over the time as a result of non-stationarity and accordingly the initial trained model is no more optimal for MI detection in evaluation session. Adapting an EEG processing unit has been used to address non-stationarity [41, 44]. An EEG processing unit in a BCI system is responsible

for feature extraction and classification (please see Section 2.1). Therefore, adaptation is applied on either feature extraction or classification part. In Section 7.2 some of these studies are reviewed. An adaptive method needs to be fast enough to be practically applicable in online BCI scenarios.

In recent years, Extreme Learning Machine (ELM) algorithm has attracted lots of attention [180–182]. ELM proposed by Huang et al. [180] is an interesting technique with faster learning speed and better generalization in comparison with traditional feed-forward neural networks. ELM has been used in different studies. Incremental learning of ELM was previously studied in [183, 184]. They applied it for human face and action recognition. Another study used ELM to automatically detect epileptic seizures [182]. It has been shown that ELM is an effective tool for those applications. Fast training time of ELM makes it suitable for online analysis. Most BCI systems in real applications are real time. Therefore, it may be assumed that ELM method can be used for MI detection in EEG-based MI-BCI systems. To the best of our knowledge this study used ELM for the first time in BCI applications. ELM method is briefly reviewed in Section 7.3. In this chapter we mainly focus on developing a novel adaptive ELM algorithm to address inter-session non-stationarity and thus enhance the session-to-session transfer performance. We seek to make our proposed algorithm optimal for online applications. Our proposed method, adaptive ELM (A-ELM), is fully explained in Section 7.4.

7.2 Review adaptive methods for addressing inter-session non-stationarity

Inter-session non-stationarity causes some changes in feature space distribution of EEG data. Shenoy et al. in [41] showed that there is a statistical dif-

ference between the calibration and online evaluation session and interpreted it as a shift in feature space. Generally speaking, adaptation has been the most commonly used method to address non-stationarity [39, 170–173, 185–187]. The adaptive methods proposed to address EEG non-stationarity can be categorized into two major groups. One group of studies focused on adaptive feature extraction methods (please see Section 6.2.2). Another group developed adaptive classification methods. Since the focus of this chapter is on developing a novel adaptive classifier, this section only reviews some of the proposed adaptive classification methods that addressed inter-session non-stationarity.

Adaptive classifiers are evolved to overcome the changes in data from one session to another session or even within a single session. Bias adaptation for LDA classifier was proposed in [41] and compared with other adaptive techniques. The results suggested that their proposed adaptive classifier overcomes the shift in distribution of the data.

Adaptive LDA classifier was also applied to a fully online BCI system [187]. The initial LDA classifier was adaptively updated through adaptive estimation of the information matrix. The results showed an improvement in performance of the subjects from one session to another session. In [42] Kalman adaptive LDA and adaptive information matrix QDA was studied. It was shown that both of these continuously adaptive classifiers outperformed discontinuously adaptive ones.

An unsupervised adaptive method of the LDA classifier was also proposed in [33]. It was shown that their proposed method was effective for online BCI system. Ang et al. in [118] used a novel online adaptive and semi-supervised learning method using NBPW classifier. They used FBCSP for feature extraction [118] and reported that the averaged Kappa value enhanced using their proposed adaptive NBPW classifier.

Liyanage at al. in [188] proposed a dynamically weighted ensemble classification method. Multiple classifiers were trained based on clustered features. Their results showed that the SVM classifier with dynamic weighting significantly outperforms the normal SVM classifier. They showed that clustering was helpful in solving inter-session non-staionarity.

In conclusion, although different adaptive classifier has been proposed over the past years, the best adaptive classification algorithm in real online BCI applications was bias adaptation so far [41, 42]. The method can fast adjust to recent recorded data from evaluation session. However, many researchers are still looking for better adaptive classification methods.

7.3 Brief review of ELMs

An ELM is a Single-Hidden Layer Feed-Forward Neural Network (SLFN) [180]. Figure 7.1 shows a simplified structure of ELM used in our study. The number of output nodes are chosen according to the number of classes. The hidden layer nodes parameters are selected randomly while the output weights are determined analytically. Since the weights of the hidden layer are assigned randomly, the learning process is performed at an extremely fast speed. In fact, ELM converts a learning problem into a linear system whose output weights can be determined through inverse operation of hidden layer weight matrices.

For n_t given training feature set $(\mathbf{x}_j, \mathbf{y}_j)$ where $\mathbf{x}_j \in \mathbb{R}^{\tilde{m}}$, $\mathbf{y}_j \in \mathbb{R}^2$, and $\tilde{m}$ denotes the number of selected features, the output of a standard ELM with activation function $g(x)$ and $\tilde{N}$ hidden nodes is calculated as follows:

$$\sum_{i=1}^{\tilde{N}} \beta_i g(\mathbf{w}_i \mathbf{x}_j + \mathbf{b}_i) = \mathbf{y}_i, \ j = \{1, 2, \cdots, n_t\} \tag{7.1}$$

where $(\mathbf{w}_i, \mathbf{b}_i)$ are randomly assigned weight and bias of the ith hidden node

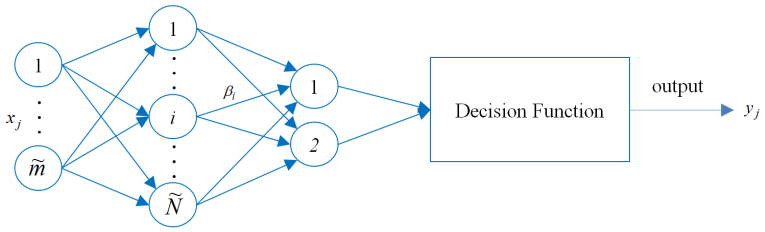


Figure 7.1: A simplified ELM structure for a two class problem. Given a training set (x_j,y_j), β_i is the output weight vector of hidden node i. The number of input nodes $\tilde{m}$ is the number of CSP features used for training an ELM and $\tilde{N}$ is the number of hidden nodes chosen arbitrary.

and $\boldsymbol{\beta}_i$ is the output weight. Eq. 7.1 can be compactly written as:

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{Y} \tag{7.2}$$

In order to find the output weights $\boldsymbol{\beta}$, the following least-square fitting is solved:

$$\min_{\boldsymbol{\beta}} \|\mathbf{H}\boldsymbol{\beta} - \mathbf{Y}\| \tag{7.3}$$

the optimal output weights $\hat{\boldsymbol{\beta}}$ is then obtained:

$$\hat{\boldsymbol{\beta}} = \mathbf{H}^{\dagger}\mathbf{Y} = \boldsymbol{\Psi}^{-1}\mathbf{H}^{T}\mathbf{Y} \tag{7.4}$$

$$\boldsymbol{\Psi} = (\mathbf{H}^{T}\mathbf{Y})^{-1} \tag{7.5}$$

where $\mathbf{H}^{\dagger}$ denotes the Moore-Penrose pseudo-inverse of the hidden-layer output matrix $\mathbf{H}$ and ^{T} denotes the transpose operator. By computing $\hat{\boldsymbol{\beta}}$ from the calibration data, the ELM classifier is trained for detecting MI in evaluation session.

7.3.1 ELM versus SVM

Both ELM and SVM methods have the same general idea of mapping the original space into a feature space in which the problem can be solved by

a linear approach. However, they have some differences in theory. As an example bias b in SVM should not be zero while for an ELM bias b is not needed [181]. In comparison with SVM [181], ELM has superior computational time. The ELM superiority increases remarkably by augmenting the number of training samples. Nevertheless, ELM has less generalization ability comparing to SVM, especially in small training sample size. Finally, the performance of SVM and ELM have been shown [181] to be quite comparable in different studies.

7.4 Adaptive ELM (A-ELM)

In order to adapt the ELM classifier, the output weights of the trained ELM needs to be updated according to the recently recorded EEG data. To derive the update rule for the output weight, a first chunk of data from evaluation session is used and the initial minimization error problem is updated to [183]:

$$\min_{\boldsymbol{\beta}} \| \begin{bmatrix} \mathbf{H}_0 \\ \mathbf{H}_1 \end{bmatrix} \boldsymbol{\beta} - \begin{bmatrix} \mathbf{Y}_0 \\ \mathbf{Y}_1 \end{bmatrix} \| \tag{7.6}$$

where $\mathbf{H}_0$ is the output of hidden layer using the training data from calibration session and $\mathbf{Y}_0$ is training data labels. Accordingly, $\mathbf{H}_1$ is the output of hidden layer for the first chunk of evaluation data and $\mathbf{Y}_1$ is their corresponding labels.

Therefore, the new output weight is calculated based on least-square minimization as follows:

$$\hat{\boldsymbol{\beta}}^{(1)} = \boldsymbol{\Psi}_1^{-1} \begin{bmatrix} \mathbf{H}_0 \\ \mathbf{H}_1 \end{bmatrix}^T \begin{bmatrix} \mathbf{Y}_0 \\ \mathbf{Y}_1 \end{bmatrix} \tag{7.7}$$

where it can be expanded to:

$$\hat{\boldsymbol{\beta}}^{(1)} = \boldsymbol{\Psi}_1^{-1}(\boldsymbol{\Psi}_0\boldsymbol{\Psi}_0^{-1}\mathbf{H}_0^T\mathbf{Y}_0) = \boldsymbol{\Psi}_1^{-1}(\boldsymbol{\Psi}_0\hat{\boldsymbol{\beta}}^{(1)} + \mathbf{H}_1^T\mathbf{Y}_1) \tag{7.8}$$

where $\boldsymbol{\Psi}_1$ is calculated as follows:

$$\boldsymbol{\Psi}_1 = \begin{bmatrix}\mathbf{H}_0\\ \mathbf{H}_1\end{bmatrix}^T \begin{bmatrix}\mathbf{H}_0\\ \mathbf{H}_1\end{bmatrix} = \boldsymbol{\Psi}_0 + \mathbf{H}_1^T\mathbf{H}_0 \tag{7.9}$$

Substituting $\boldsymbol{\Psi}_1$ from 7.9, we have:

$$\hat{\boldsymbol{\beta}}^{(1)} = \hat{\boldsymbol{\beta}}^{(0)} + \boldsymbol{\Psi}_1^{-1}\mathbf{H}_1^T(\mathbf{Y}_1 - \mathbf{H}_1\hat{\beta}^{(0)}) \tag{7.10}$$

In order to calculate the inverse of $\boldsymbol{\Psi}_1$ we used the matrix inversion lemma. This lemma states that for a given matrix $\mathbf{A} = (\mathbf{B} + \mathbf{UDV})$, its inverse is determined by:

$$\mathbf{A}^{-1} = \mathbf{B}^{-1} - \mathbf{B}^{-1}\mathbf{U}(\mathbf{D}^{-1} + \mathbf{VB}^{-1}\mathbf{U})^{-1}\mathbf{VB}^{-1}$$

We use this lemma to get the inverse of $\boldsymbol{\Psi}_1$ defined in 7.9. Finally, the recursive formulation for updating the output weights can be defined as follows:

$$\hat{\boldsymbol{\beta}}^{(k+1)} = \hat{\boldsymbol{\beta}}^{(k)} + \boldsymbol{\Psi}_{k+1}^{-1}\mathbf{H}_{k+1}^T(\mathbf{Y}_{k+1} - \mathbf{H}_{k+1}\hat{\boldsymbol{\beta}}^{(k)}) \tag{7.11}$$

$$\boldsymbol{\Psi}_{k+1} = \boldsymbol{\Psi}_k + \mathbf{H}_{k+1}^T\mathbf{H}_{k+1} \tag{7.12}$$

$$\boldsymbol{\Psi}_{k+1}^{-1} = \boldsymbol{\Psi}_k^{-1} + \boldsymbol{\Psi}_k^{-1}\mathbf{H}_{k+1}^T[\mathbf{I} + \mathbf{H}_{k+1}\boldsymbol{\Psi}_k^{-1}\mathbf{H}_{k+1}^T]\mathbf{H}_{k+1}\boldsymbol{\Psi}_k^{-1} \tag{7.13}$$

The adaptive ELM algorithm explained above has two steps: initialization and adaptation. During first step or initialization the training EEG data recorded during calibration session is used. The data of the evaluation

session is used during second step or adaptation. Overall, A-ELM algorithm is explained in Algorithm 2.

Algorithm 2 Adaptive extreme learning machine (A-ELM)

Part I: Initialization
Given a training feature sample $(\mathbf{x}_i, \mathbf{y}_i)$, an activation function $g : \mathbb{R} \to \mathbb{R}$, and $\tilde{N}$ number of hidden nodes:

1: Randomly assign the weights w_i and bias b_i of the hidden layer.
2: Compute the output vector $\mathbf{H}$ of the hidden layer (Eq. 7.1).
3: Calculate the initial output weight $\hat{\boldsymbol{\beta}}^{(0)}$ according to Eq. 7.4.

Part II: Adaptation

1: Select a recent chunk of EEG data from the evaluation session.
2: Estimate the labels of the trials in the selected chunk based on the current settings of the ELM.
3: Update the output weights according to Eq. 7.11 to Eq. 7.13.
4: Repeat steps (1-3) of ***Part II*** until all the labels of evaluation data are estimated.

The number of input nodes $\tilde{m}$ is the number of features selected for training the classifier. The number of hidden nodes $\tilde{N}$ is selected based on CV accuracies of the train data for different value of the hidden nodes. The number of trials used in adaptation can be selected arbitrary. Since the final goal of the proposed algorithm is to be used in online scenario, it is preferred to have a smooth change in classifier. Every 5 to 10 trials of evaluation data can be used for adapting the classifier. The activation function g can be sigmoid, sine or hard-limit function.

7.5 Experimental setup

In this chapter, the EEG data collected from 17 healthy subjects were used. Two subjects were left-handed and rest were right-handed. The right (left) handed subjects were asked to perform right (left) hand MI. All the subjects were asked for ethics and approval and consent. EEG signals were collected using the Nuamps EEG acquisition hardware with unipolar Ag/AgCl electrodes channels, digitally sampled at 250 Hz with a resolution of 22 bits for voltage ranges of ± 130 mV. EEG recordings from all 27 channels were band

pass filtered from 0.05 to 40 Hz by the acquisition hardware. Prior to the experiments, the subjects were instructed to minimize any physical movement and eye blinking throughout the EEG recording process.

The EEG data from each subject was collected in two sessions each on a separate day. On the first day, a calibration session was recorded that contained two non-feedback runs. The evaluation session was recorded on another day and had three non-feedback runs. During these sessions, the subjects were instructed to perform kinaesthetic MI of their chosen hand versus idle state right after a visual cues displayed on the computer screen in each trial. Each session comprised of 40 trials of MI and 40 trials of idle state and lasted about 16 minutes. Each trial comprised a preparatory segment of 2s, the presentation of the visual cue for 4 seconds, and a rest segment of at least 6s. Each trial lasted approximately 12 seconds, and a break period of at least 2 minutes was given after each run of EEG recording.

7.6 EEG data processing

7.6.1 Pre-processing

The EEG data from 0.5 to 2.5 s after providing the cue was extracted and used for feature extraction. The selected time segment was previously suggested by the winner of the BCI competition IV dataset IIa [117]. The extracted EEG data from all 27 channels was then band-pass filtered between 8 to 30 Hz using a fifth-order Butterworth filter. This frequency band includes both alpha and beta frequency bands which are mainly involved in performing MI. Subsequently, the band-pass filtered data were spatially filtered to select the most discriminative features. The first and last two CSP filters, $m = 2$, were chosen. Hence, totally $\tilde{m} = 4$ features were selected to be applied to the classifier.

7.6.2 Evaluation of session-to-session transfer accuracy

The recorded EEG signal in the calibration session was preprocessed and 2 pairs of features were then extracted from calibration data and used for training the classifier. The calibration-to-evaluation session performances of the subjects were then evaluated by SVM, ELM and A-ELM classifiers. Figure 7.2 shows the flow diagram of our analysis. The proposed A-ELM algorithm updates the initial classifier sequentially according to a newly selected chunk of EEG data.

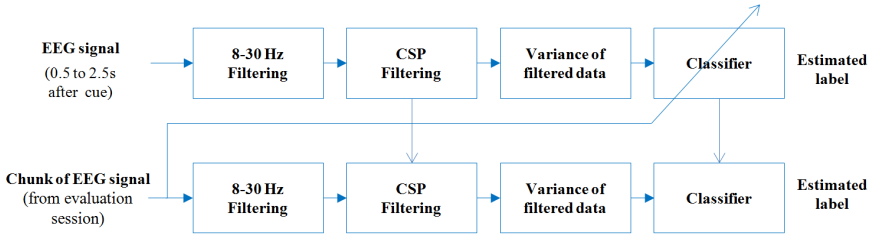


Figure 7.2: Methodology of our analysis.

Here in this chapter we proposed three methods for adaptively updating of A-ELM: 1) *Balanced*: the selected chunk of data contained equal number of trials from each class, 2) *Unbalanced*: the selected chunk contained unbalanced number of trials from each class which means that the trials were in the same order as they were recorded, and 3) *Accumulated*: the selected previous chunks were also used for adaptation. The A-ELM methods were then applied on the data set collected from 17 healthy subjects explained in Section 7.5.

7.7 Results

7.7.1 ELM initialization

As stated earlier, the number of ELM input nodes is the same as the number of features $\tilde{m} = 4$. In order to choose the best number of hidden nodes, we

calculated the 5×5-fold CV accuracies of subjects in MI calibration session using various number of hidden nodes. Figure 7.3 shows how the averaged CV accuracy over all 17 subjects varies by choosing different number of hidden nodes.

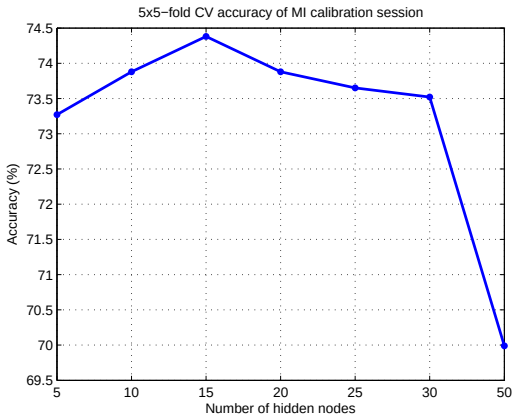


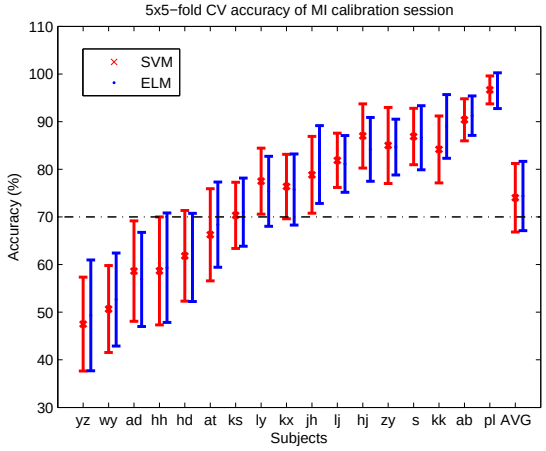
Figure 7.3: Choosing the best number of hidden nodes. Each dot (·) represents the averaged 5×5-fold CV accuracies of MI calibration session collected from 17 healthy subjects for different number of hidden nodes.

The best number of hidden nodes was $\tilde{N} = 15$, and thus in all analysis provided in the following sections the number of ELM hidden nodes is fixed at 15. However, as can be seen in Figure 7.3, changing the number of hidden nodes $\tilde{N}$ does not have a great impact on averaged CV accuracies.

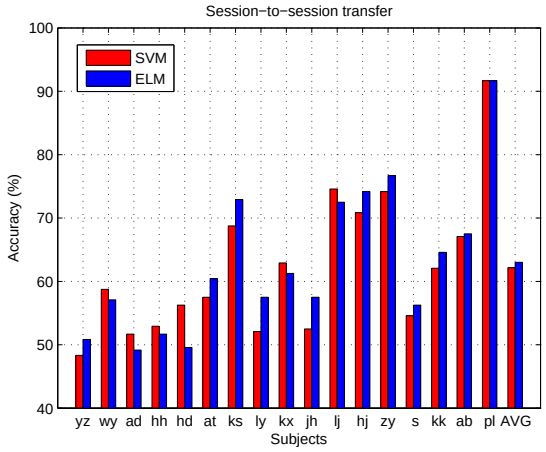
7.7.2 Baseline classification results

The performance of subjects in MI calibration session is shown in Figure 7.4 (a). The 5×5-fold CV accuracies of the calibration session using SVM and ELM classifiers were compared. The average performance of subjects using ELM classifier was 74.38%± 14.30, while the averaged performance of SVM classifier was 74.03%±14.54. A paired sample t-test showed no significant difference between SVM and ELM classifier $t(16)$=-0.775, emphp=0.449. Comparing the total training time of classifier for all 17 subjects showed

ELM training was faster (86 s) than SVM training (86.39 s).



(a)



(b)

Figure 7.4: Baseline classification accuracy using ELM and SVM classifiers. The subjects are sorted according to their ELM accuracies in calibration session. (a) Comparing 5×5-fold CV accuracies in MI calibration session. Errorbars indicate the standard deviation over CV folds; (b) Comparing ELM and SVM accuracies in calibration-to-evaluation session transfer.

The performance of ELM and SVM classifiers in session-to-session transfer are shown in Figure 7.4 (b). There was also no significant difference between SVM and ELM classifier $t(16)$=-1.109, p=0.284 in session-to-session transfer. Some subjects such as *jh*, *kk*, *s* and *ly* had calibration accuracy above 70%, while in session-to-session transfer their performance dropped

below 70%. Subjects with poor performance in calibration session had also poor performance in evaluation session.

7.7.3 Visualizing inter-session non-stationarity

As mentioned earlier, the distribution of the features in calibration session and evaluation session may be different due to non-staionarity of EEG data. Figure 7.5 visualizes the possible differences between the sessions.

Figure 7.5 (a)-(d) are the results of four selected subjects (*jh*, *kk*, *s*, *ly*) with poor session-to-session transfer performance. Comparing their first CSP filter and pattern of calibration and evaluation session showed that all four subjects performed MI task similarly in both sessions. However, the feature distribution changed from one session to another, such inter-session non-stationarity led to inseparable features in evaluation session.

Figure 7.5 (g)-(h) shows the results for four selected subjects with high performance. The two dimensional feature space of high performance subjects was also changed, but the change was not big enough to affect the separability of features.

7.7.4 Session-to-session transfer results

In this section, session-to-session classification accuracy of subjects using A-ELM with three different sequential learning methods (i.e., *Balanced*, *Unbalanced*, *Accumulated*) explained in Section 7.6.2 is evaluated. The results are shown in Figure 7.6. The first row compares the non adaptive ELM with A-ELM. The accuracies lied above the diagonal showed the better accuracy for A-ELM. The second row compares the three proposed A-ELM methods.

Using *Balanced* A-ELM method resulted in higher accuracy than non-adaptive ELM for 15 out of 17 subjects. This number increased to 16 and 17 subjects when using *Unbalanced* and *Accumulated* A-ELM. As can be seen,

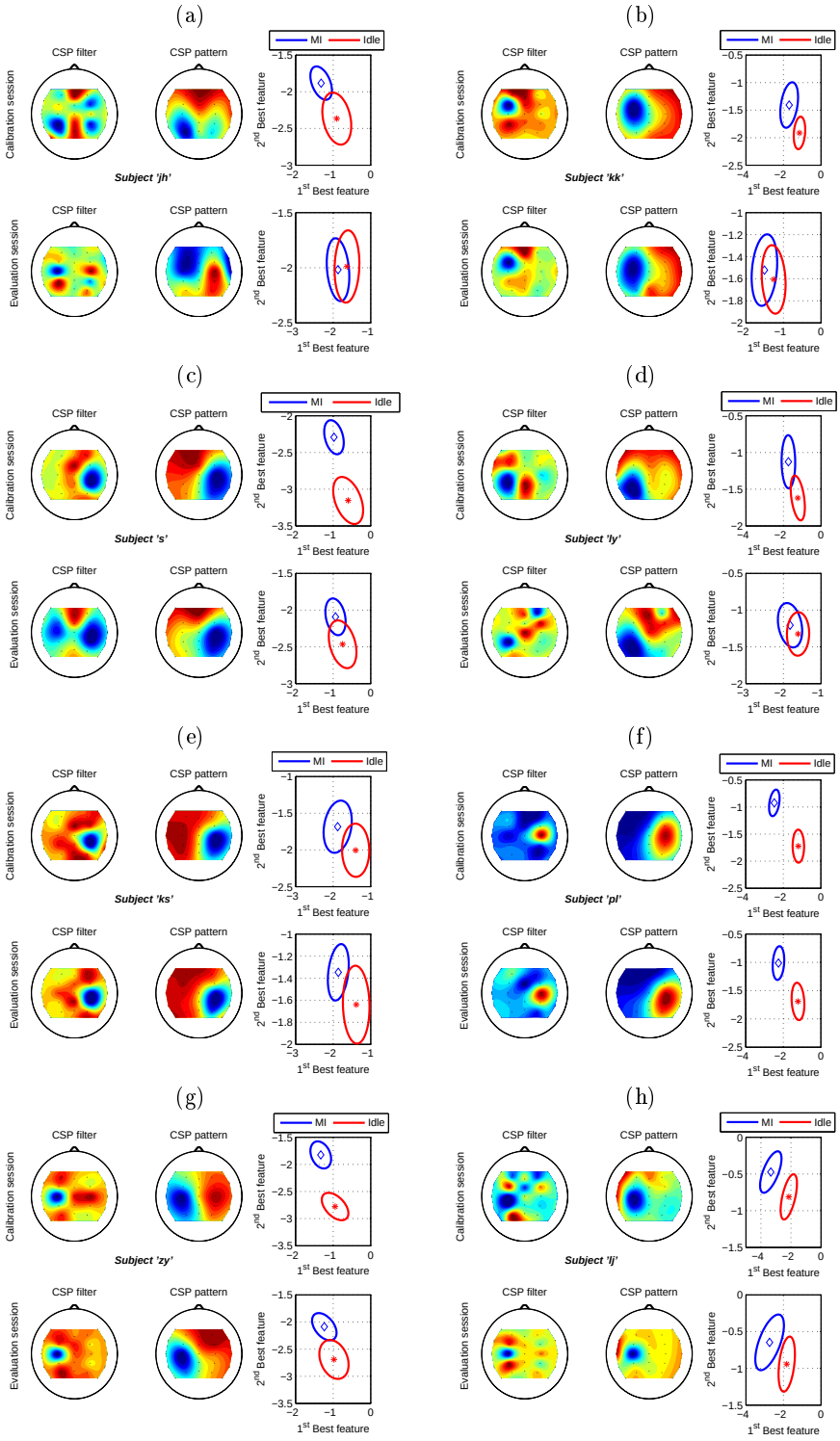


Figure 7.5: Comparing calibration and evaluation session for selected subjects: '*jh*' (a), '*kk*' (b), '*s*' (c), '*ly*' (d), '*ks*' (e), '*pl*' (f), '*zy*' (g) and '*lj*' (h). Column 1 and 2 are first CSP filter and pattern. Column 3 shows the two dimensional feature space. The ellipsoid are distributions of features. The color-scale is in arbitrary units.

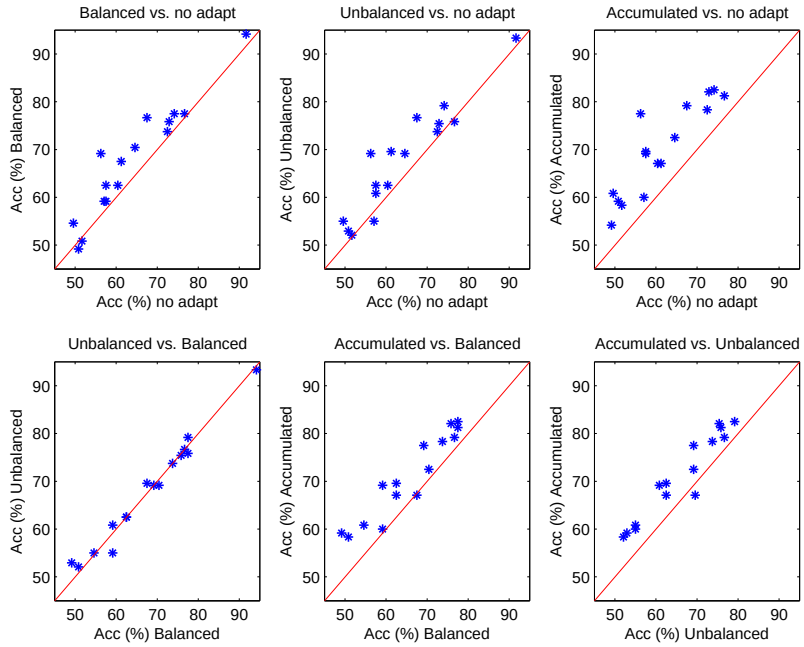
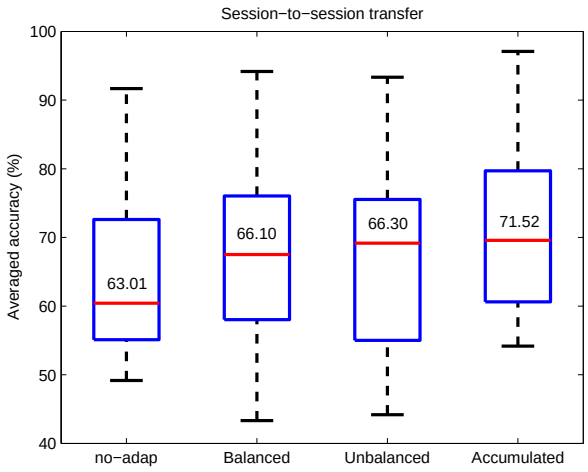


Figure 7.6: Comparison of various A-ELM methods (*Balanced*, *Unbalanced*, *Accumulated*) with non adaptive ELM (no adapt) in session-to-session transfer. The accuracy of subjects are shown by asterisks (*). Each subplot shows a scatter plot of subjects' accuracies for two different methods.

the results of *Balanced* and *Unbalanced* A-ELM methods are quite similar. Among the three proposed A-ELM methods, *Accumulated* A-ELM showed higher accuracies for almost all the subjects.

The averaged session-to-session transfer accuracies of the subjects are shown in Figure 7.7(a). The averaged performance of all three A-ELM methods was higher than the averaged baseline accuracy and *Accumulated* A-ELM method resulted in highest averaged performance (71.52%). A paired sample t-test was conducted and the corresponding p-values are listed in Figure 7.7(b). Conducting a sampled paired t-test showed that the average accuracy of subjects using ELM without adaptation was significantly lower than that of all three A-ELM methods. However, there was no significant difference in balanced or unbalanced selection of trials (p=0.648).



(a)

Method	*Balanced*	*Unbalanced*	*Accumulated*
No adaptation	0.009	0.006	<0.001
Balanced	-	0.648	<0.001
Unbalanced	0.648	-	<0.001

(b)

Figure 7.7: Comparing the average performance of 17 healthy subjects using non-adaptive ELM and A-ELM (*Balanced*, *Unbalanced*, *Accumulated*). (a) Box plot (b) p-values derived from paired t-test.

To evaluate the effect of increasing the number of samples in adapting the classifier, *Accumulated* A-ELM was proposed. In this method, the selected chunks of evaluation data were accumulated and used for adaptation. This increased the average accuracy to 71.52% which was significantly higher (p <0.001) than non-adaptive ELM, *Balanced* and *Unbalanced* methods.

7.8 Discussion

The results of ELM and SVM classifiers in Figure 7.4 showed no significant difference between these two classifiers. This is consistent with the results reported in previous works [181]. However, as we expected the training time of ELM was less than SVM. Although the time difference was not very

remarkable in our study, it might speed up the training time for subject independent classifiers used in BCI.

Visualizing the feature space in session-to-session transfer helps us better interpret the performance variation of the subjects. Figure 7.5(a)-(d) shows the selected subjects performing MI similarly in both calibration and evaluation session, nevertheless their feature distribution was different. This can be possibly one of the reasons of their poor performance. In fact, having such shifts in feature space makes the initial classifier to be suboptimal for the evaluation session. On the other hand, we can hardly see such changes from calibration into evaluation in Figure 7.5(e)-(h), and thus as we expected the selected subjects had high performance accuracy in session-to-session transfer. Non-stationarity of EEG is one of the possible reasons of poor performance.

A-ELM method was proposed in this chapter to alleviate the inter-session non-stationarity. Three different learning algorithms were used for adapting the classifier. The results showed that adaptation significantly enhances the performance of the subjects. Using trials *Unbalanced* with their original recording order resulted in slightly better averaged accuracy than using *Balanced* trials from each class. This can be possibly attributed to the time correlation of trials. Using *Accumulated* ELM significantly enhanced the accuracies in comparison with other A-ELM methods. This implies that using more data from evaluation session was helpful to overcome the differences between two sessions.

7.9 Summary

This chapter aimed to propose a method for improving the session-to-session transfer performance of MI-BCI. Non-stationarity of EEG signal decreases

session-to-session transfer performance of the users, so adaptive extreme learning machine (A-ELM) was applied to compensate such deterioration. A-ELM used limited number of EEG trials from the evaluation session (i.e., at least one chunk of EEG data) to adaptively update the initial ELM classifier. The results showed that A-ELM had significantly better performance in comparison with baseline classifiers (non-adaptive ELM and SVM). The results also suggested that using more data from evaluation session for adapting the classifier significantly improves the performance of the users. In contrast to most adaptive methods based on updating the features, A-ELM does not need balanced data for adaptation. In fact, there was no significant difference in performance of the users when balanced or unbalanced data from the evaluation session were selected for updating the classifier. In conclusion, ELM can be considered as one of the appropriate solutions for online BCI systems due to its fast learning process and acceptable performance. Fast training of ELM classifier makes it more suitable in presence of large training data such as subject independent classifier. Therefore, we can benefit from using ELM in BCI applications.

Chapter 8

Conclusions and future works

8.1 Summary and conclusion

This thesis mainly focuses on improving the performance of EEG-based MI-BCI. The BCI performance shows how well a subject can control a BCI system. Two main reasons for poor performance in EEG-based MI-BCI are: 1) BCI deficiency, and 2) EEG non-stationarity. Addressing these issues may yield a more practical BCI system that can be controlled by all subjects. To do so, we can help subjects with BCI deficiency improve their performance and also develop adaptive algorithms. The subjects with BCI deficiency cannot modulate their brain signals. To help these subjects improve their performance, first we should identify them. Therefore, two performance predictors were proposed to detect BCI deficient subjects, and then a novel experiment was designed to help subjects enhance their performance. Subsequently, two adaptive algorithms were developed to address inter-session non-stationarity and to improve the BCI performance.

In Chapter 3, a novel neurophysiological coefficient from pre-cue EEG rhythms was proposed to predict the MI-BCI performance. The state of the brain preceding a task affects the subsequent task outcome. Thus, pre-cue EEG was used to predict the performance of subjects. Attention level is one of the indicators that defines the state of the brain. Therefore, quantifying the attention in the pre-cue time segment led us to propose a novel coefficient for predicting MI-BCI performance. The proposed neurophysiological coef-

ficient which includes brain topographic information, is one of the possible quantification of attention level. In fact, the proposed coefficient captures both spectral and spatial information and yields a better representation of the brain state. The experimental results showed significant positive correlation between the proposed neurophysiological performance predictor and the BCI classification accuracy of the subjects. The results suggested that increases in performance were correlated with increases in subject's attention level. Therefore, we can conclude that higher performance subjects have a higher attention level and vice versa.

The second proposed predictor in Chapter 4 used the SSD method to extract the alpha neural oscillations by taking into account the spatial structure of noise. The role of alpha band on BCI performance has been investigated in several studies. However, the extracted alpha activity contained background noise. The SSD-based neurophysiological performance predictor was proposed to elicit spatial-spectral components from EEG data with high signal-to-noise ratio. The experimental results on a big dataset showed that the proposed predictor based on resting state EEG in eyes open condition was significantly correlated to the MI-BCI performance of subjects. This result suggested that increases in resting sate alpha were correlated to increases in BCI-performance. On top of that, although resting state alpha power in the eyes closed condition was higher than the eyes open condition, the later had higher correlation with BCI performance than the former. These findings suggested that resting state alpha in the eyes open condition can be used to predict the performance of the subjects prior to the experiment. Group level analysis showed that the proposed predictor in pre-cue EEG was also significantly correlated to BCI performance. This finding is promising enough to use the predictor for single trial performance prediction. The results also revealed that incorporating theta and beta frequency

bands enhanced the correlation values. In conclusion, the results of the proposed SSD-based predictor outperformed the spectral-based predictor which was previously proposed in the literature. This can be mainly because the extracted components have enhanced signal-to-noise ratio in comparison to normal extracted alpha band-power and thus led to a better prediction of the BCI performance.

After predicting the performance of the subjects, we seek a method to help BCI deficient subjects control a MI-BCI. Thus, in Chapter 5 a novel experimental design was proposed to improve the MI-BCI performance. Since there was significant correlation between resting state alpha band power and BCI performance, our designed experiment aimed to enhance the resting state alpha through several NFT sessions using the proposed SSD-based predictor. To the best of our knowledge, no experiment has been designed to enhance the BCI performance of the subjects through alpha NFT. Thirteen healthy subjects participated in our study and randomly assigned to experimental and control group. In the proposed experiment, the BCI performance of the experimental group was evaluated before and after the NFT sessions. The results showed that the BCI performance of these subjects significantly improved after 12 sessions of NFT. However, the BCI performance of the control group showed no significant improvement. Having small sample size may be considered as one of the limitations of this study. Having more subjects can lead to having more conclusive results. The results of NFT sessions on resting state alpha revealed that the average resting state alpha was significantly improved by NFT.

Developing adaptive algorithms is another strategy for enhancing the EEG-based MI-BCI performance. In this thesis, an adaptive feature extraction method and an adaptive classifier were proposed to address inter-session non-stationarity that causes changes in feature distribution and thus

deteriorates the performance. In view of similarities between PM and MI brain patterns, PM showed to be a plausible method for calibrating MI-BCI. Nonetheless, the difference between the PM calibration session and MI evaluation session needs to be addressed. Accordingly, in Chapter 6 we proposed a novel method to adaptively update the features extracted using FBCSP to diminish features drift across the sessions. Our proposed algorithm updates the CSP matrix of each band in FBCSP with the recent recorded batch of EEG data. In presence of big differences between the feature distribution of the calibration session and the current session, KL distance weighting gives more weight to the covariance matrix of recent recorded data. The results showed that MI-BCI, when calibrated by PM, had higher averaged performance than calibrated by MI. Thus, due to the applicability of the proposed method for online scenario, it may be inferred that calibrating MI-BCI with PM facilitates data collection and results in better online performance.

The second algorithm proposed to address inter-session non-stationarity in Chapter 7 was A-ELM. Since the distribution of the extracted CSP features from calibration and evaluation session was different, we proposed the A-ELM to update the baseline ELM classifier using limited number of EEG trials recorded in the evaluation session. The experimental results showed that A-ELM significantly improved the performance of the subjects in comparison with baseline ELM and state-of-the-art classifier, SVM. Furthermore, using accumulated EEG data from the evaluation session for adapting ELM significantly increased the averaged performance in comparison with using only limited number of data. The experimental results showed training ELM in comparison with the state-of-the-art SVM was slightly faster.

In conclusion, the fundamental purpose of the proposed methods summarized above was to come up with a more practical BCI system for real world applications. One of the main prospects of the performance predictors

is to decide whether the stroke patients could benefit from BCI-based therapy. Applying our proposed predictors over the resting state can predict the performance of the subject, thus those with poor estimated accuracy may not follow BCI therapy. Our proposed predictor can literally replace the typical screening session for categorizing the patients.

8.2 Future works

The fundamental purpose of the proposed methods in this thesis was to come up with a more practical BCI system. In future, the work presented in this thesis can be potentially extended to be used in real world applications. In the following, some of these extensions are briefly explained. The proposed performance predictors are applicable for both normal subjects and patients. As an example, one of the main prospects of the performance predictors is to decide whether the patients can benefit from the BCI-based therapy. The performance of the patients can be predicted based on a short recording of resting state EEG data. Therefore, those patients with poor estimated accuracies may follow other interventions rather than the BCI-based therapy. The resting state in comparison with the BCI screening session is much shorter. Thus the proposed predictor can literally replace the typical screening session for categorizing the patients in the future. The proposed predictors can be also used for detecting healthy subjects with BCI deficiency. The detected poor performance subjects may follow the NFT sessions to enhance their BCI performance. Moreover, in future experiments, subjects may be instructed to be more attentive prior to a cue to improve their BCI performance. In addition, a new experiment can be designed to specifically help poor performance subjects. The new designed experiment may incorporate a varying preparation time preceding a cue to help the subjects reach a predefined at-

tention level and thus better control a BCI system. Another possible design may include a small stroop test to help the subjects reach a predefined attention level in the middle of the experiment. These new experimental designs can make MI-BCI systems to be applicable for all users. Furthermore, the proposed adaptive algorithms can be applied in online BCI systems to tackle non-stationarity and thus improve the MI-BCI performance. The proposed adaptive classifier can be also extended to a subject independent classifier. The calibration session, which is usually required in most BCI experiments to train a subject specific model, is quite time consuming and a tedious step especially for patients. A subject independent classifier can be used as an alternative of a subject dependent classifier to remove the calibration session. The proposed adaptive ELM classifier, which has a fast training time in comparison with the other state-of-the-art classifiers, can be trained by the calibration data, which collected from a large number of subjects, to build a subject independent classifier.

Bibliography

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.

[2] N. Birbaumer, "Brain-computer interface research: Coming of age," *Clinical Neurophysiology*, vol. 117, no. 3, pp. 479–483, 2006.

[3] G. Dornhege, J. d. R. Millan, T. Hinterberger, D. J. McFarland, and K.-R. Muller, *Toward Brain-Computer Interfacing*. MIT press, 2007.

[4] B. Graimann, B. Allison, and G. Pfurtscheller, *Brain-Computer Interfaces: A Gentle Introduction*, ser. The Frontiers Collection. Springer Berlin Heidelberg, 2010, ch. 1, pp. 1–27.

[5] J.-J. Vidal, "Toward direct brain-computer communication," *Annual Review of Biophysics and Bioengineering*, vol. 2, no. 1, pp. 157–180, 1973.

[6] N. Birbaumer, N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey, A. Kubler, J. Perelmouter, E. Taub, and H. Flor, "A spelling device for the paralysed," *Nature*, vol. 398, no. 6725, pp. 297–298, 1999.

[7] N. Birbaumer and L. G. Cohen, "Brain-computer interfaces: communication and restoration of movement in paralysis," *The Journal of Physiology*, vol. 579, no. 3, pp. 621–636, 2007.

[8] B. Blankertz, M. Krauledat, G. Dornhege, J. Williamson, R. Murray-Smith, and K. Müller, *A Note on Brain Actuated Spelling with the*

Berlin Brain-Computer Interface. Springer Berlin / Heidelberg, 2007, vol. 4555, pp. 759–768.

[9] V. Kaiser, A. Kreilinger, G. R. Muller-Putz, and C. Neuper, "First steps towards a motor imagery based stroke BCI new strategy to set up a classifier," *Frontiers in Neuroscience*, vol. 5, pp. 1–10, Jul 2011.

[10] N. Birbaumer, A. R. Murguialday, and L. Cohen, "Brain-computer interface in paralysis," *Current Opinion in Neurology*, vol. 21, no. 6, pp. 634–638, Dec. 2008.

[11] B. Várkuti, C. Guan, Y. Pan, K. S. Phua, K. K. Ang, C. W. K. Kuah, K. Chua, B. T. Ang, N. Birbaumer, and R. Sitaram, "Resting state changes in functional connectivity correlate with movement recovery for BCI and robot-assisted upper-extremity training after stroke," *Neurorehabilitation and Neural Repair*, vol. 27, no. 1, pp. 53–62, 2013.

[12] B. Rebsamen, G. Cuntai, Z. Haihong, W. Chuanchu, T. Cheeleong, M. H. Ang, and E. Burdet, "A brain controlled wheelchair to navigate in familiar environments," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 6, pp. 590–598, 2010.

[13] B. Rebsamen, E. Burdet, C. Guan, Z. Haihong, T. Chee Leong, Z. Qiang, M. Ang, and C. Laugier, "A brain-controlled wheelchair based on P300 and path guidance," in *Annual International Conference on Biomedical Robotics and Biomechatronics*, 2006, pp. 1101–1106.

[14] K. Tanaka, K. Matsunaga, and H. O. Wang, "Electroencephalogram-based control of an electric wheelchair," *Robotics, IEEE Transactions on*, vol. 21, no. 4, pp. 762–766, 2005.

[15] K. K. Ang, C. Guan, K. S. G. Chua, B. T. Ang, C. W. K. Kuah, C. Wang, K. S. Phua, Z. Y. Chin, and H. Zhang, "A large clinical

study on the ability of stroke patients to use an EEG-based motor imagery brain-computer interface," *Clinical EEG and Neuroscience*, vol. 42, no. 4, pp. 253–258, 2011.

[16] K. K. Ang, K. S. G. Chua, K. S. Phua, C. Wang, Z. Y. Chin, C. W. K. Kuah, W. Low, and C. Guan, "A randomized controlled trial of EEG-based motor imagery brain-computer interface robotic rehabilitation for stroke," *Clinical EEG and Neuroscience*, pp. 1–11, 2014.

[17] K. K. Ang, C. Guan, C. K. S. G., B. T. Ang, K. C., C. Wang, K. S. Phua, Z. Y. Chin, and H. Zhang, "Clinical study of neurorehabilitation in stroke using EEG-based motor imagery brain-computer interface with robotic feedback," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2010, pp. 5549–5552.

[18] J. N. Mak and J. R. Wolpaw, "Clinical applications of brain-computer interfaces: Current state and future prospects," *IEEE Reviews in Biomedical Engineering*, vol. 2, pp. 187–199, 2009.

[19] J. d. R. Millán, R. Rupp, G. R. Müller-Putz, R. Murray-Smith, C. Giugliemma, M. Tangermann, C. Vidaurre, F. Cincotti, A. Kübler, and R. Leeb, "Combining brain-computer interfaces and assistive technologies: state-of-the-art and challenges," *Frontiers in Neuroscience*, vol. 4, 2010.

[20] B. Blankertz, M. Tangermann, C. Vidaurre, S. Fazli, C. Sannelli, S. Haufe, C. Maeder, L. E. Ramsey, I. Sturm, G. Curio, and K. R. Mueller, "The berlin brain-computer interface: Non-medical uses of BCI technology," *Frontiers in Neuroscience*, vol. 4, 2010.

[21] J. van Erp, F. Lotte, and M. Tangermann, "Brain-computer interfaces: Beyond medical applications," *Computer*, vol. 45, no. 4, pp. 26–34, 2012.

[22] L. Bonnet, F. Lotte, and A. Lecuyer, "Two brains, one game: Design and evaluation of a multiuser BCI video game based on motor imagery," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 5, no. 2, pp. 185–198, 2013.

[23] D. Coyle, J. Principe, F. Lotte, and A. Nijholt, "Guest editorial: Brain/neuronal - computer game interfaces and interaction," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 5, no. 2, pp. 77–81, 2013.

[24] C. Neuper, M. Wörtz, and G. Pfurtscheller, *ERD/ERS patterns reflecting sensorimotor activation and deactivation.* Elsevier, 2006, vol. Volume 159, pp. 211–222.

[25] A. Bashashati, M. Fatourechi, R. K. Ward, and G. E. Birch, "A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals," *Journal of Neural Engineering*, vol. 4, no. 2, p. R32, 2007.

[26] J. R. Wolpaw, D. J. McFarland, G. W. Neat, and C. A. Forneris, "An EEG-based brain-computer interface for cursor control," *Electroencephalography and Clinical Neurophysiology*, vol. 78, no. 3, pp. 252–259, 1991.

[27] G. Pfurtscheller, C. Neuper, C. Guger, W. Harkam, H. Ramoser, A. Schlogl, B. Obermaier, and M. Pregenzer, "Current trends in Graz brain-computer interface (BCI) research," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 216–219, 2000.

[28] T. M. Vaughan, D. J. McFarland, G. Schalk, W. A. Sarnacki, D. J. Krusienski, E. W. Sellers, and J. R. Wolpaw, "The wadsworth BCI research and development program: at home with BCI," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 229–233, 2006.

[29] A. Kübler, V. K. Mushahwar, L. R. Hochberg, and J. P. Donoghue, "BCI meeting 2005-workshop on clinical issues and applications," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 131–134, 2006.

[30] A. P. kübler, F. M. Nijboer, J. M. Mellinger, T. M. B. Vaughan, H. P. Pawelzik, G. M. Schalk, D. J. P. McFarland, N. P. Birbaumer, and J. R. M. Wolpaw, "Patients with ALS can use sensorimotor rhythms to operate a brain-computer interface," *Neurology*, vol. 64, no. 10, pp. 1775–1777, 2005.

[31] F. Popescu, B. Blankertz, and K.-R. Müller, "Computational challenges for noninvasive brain computer interfaces," *IEEE Intelligent Systems*, vol. 23, no. 3, pp. 78–79, 2008.

[32] E. Vaadia and N. Birbaumer, "Grand challenges of brain computer interfaces in the years to come," *Frontiers in Neuroscience*, vol. 3, 2009.

[33] C. Vidaurre, C. Sannelli, K.-R. Müller, and B. Blankertz, "Co-adaptive calibration to improve BCI efficiency," *Journal of Neural Engineering*, vol. 8, no. 2, pp. 1–8, 2011.

[34] ——, "Machine-learning-based coadaptive calibration for brain-computer interfaces," *Neural Computation*, vol. 23, no. 3, pp. 791–816, 2010.

[35] B. Blankertz, C. Sannelli, S. Halder, E. M. Hammer, A. Kübler, K.-R. Müller, G. Curio, and T. Dickhaus, "Neurophysiological predictor of SMR-based BCI performance," *NeuroImage*, vol. 51, no. 4, pp. 1303–1309, 2010.

[36] E. M. Hammer, S. Halder, B. Blankertz, C. Sannelli, T. Dickhaus, S. Kleih, K.-R. Müller, and A. Kübler, "Psychological predictors of SMR-BCI performance," *Biological Psychology*, vol. 89, no. 1, pp. 80–86, 2012.

[37] W. Burde and B. Blankertz, "Is the locus of control of reinforcement a predictor of brain-computer interface performance?" in *International Brain-Computer Interface Workshop and Training Course*, 2006, pp. 76–77.

[38] M. Grosse-Wentrup and B. Schölkopf, "High gamma-power predicts performance in sensorimotor-rhythm brain-computer interfaces," *Journal of Neural Engineering*, vol. 9, no. 4, p. 046001, 2012.

[39] P. v. Bunau, F. C. Meinecke, and K.-R. Müller, "Stationary subspace analysis," *American Physics Letter*, 2009.

[40] M. Arvaneh, C. Guan, K. K. Ang, and C. Quek, "Optimizing spatial filters by minimizing within-class dissimilarities in electroencephalogram-based brain-computer interface," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 4, pp. 610–619, 2013.

[41] P. Shenoy, M. Krauledat, B. Blankertz, R. P. N. Rao, and K.-R. MÃŒller, "Towards adaptive classification for BCI," *Journal of Neural Engineering*, vol. 3, no. 1, pp. R13–R23, 2006.

[42] C. Vidaurre, A. Schlogl, R. Cabeza, R. Scherer, and G. Pfurtscheller, "Study of on-line adaptive discriminant analysis for EEG-based brain computer interfaces," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 3, pp. 550–556, 2007.

[43] Y. Li and C. Guan, "An extended EM algorithm for joint feature extraction and classification in brain- computer interfaces," *Neural computation*, vol. 18, no. 11, pp. 2730–2761, Nov. 2006.

[44] S. Sun and C. Zhang, "Adaptive feature extraction for eeg signal classification," *Medical & Biological Engineering & Computing*, vol. 44, pp. 931–935, 2006.

[45] A. Bamdadian, C. Guan, K. K. Ang, and J. Xu, "The predictive role of pre-cue EEG rhythms on MI-based BCI classification performance," *Journal of Neuroscience Methods*, vol. 235, pp. 138–144, 2014.

[46] A. Bamdadian, "Predicting the BCI performance of the subjects using spatio-spectral information of EEG signal," vol. (*Under preparation*).

[47] A. Bamdadian, C. Guan, K. K. Ang, and J. Xu, "Online semi-supervised learning with kl distance weighting for motor imagery-based BCI," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2012.

[48] A. Bamdadian, G. Cuntai, A. Kai Keng, and X. Jianxin, "Improving session-to-session transfer performance of motor imagery-based BCI using adaptive extreme learning machine," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2013, pp. 2188–2191.

[49] A. Bamdadian, "Extreme learning machine improves session-to-session transfer perfor- mance of EEG-based MI-BCI," vol. (*Under preparation*).

[50] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, and T. M. Vaughan, "Brain-computer interface technology: a review of the first international meeting," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 164–173, 2000.

[51] R. Leeb, F. Lee, C. Keinrath, R. Scherer, H. Bischof, and G. Pfurtscheller, "Brain-computer communication: Motivation, aim, and impact of exploring a virtual apartment," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 15, no. 4, pp. 473–482, 2007.

[52] M. Krauledat, M. Tangermann, B. Blankertz, and K.-R. Müller, "Towards zero training for brain-computer interfacing," *PLoS ONE*, vol. 3, no. 8, p. e2967, 2008.

[53] B. Blankertz, G. Dornhege, M. Krauledat, K.-R. Müller, and G. Curio, "The non-invasive berlin brain-computer interface: Fast acquisition of effective performance in untrained subjects," *NeuroImage*, vol. 37, no. 2, pp. 539–550, 2007.

[54] F. Lotte and G. Cuntai, "Learning from other subjects helps reducing brain-computer interface calibration time," in *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, pp. 614–617.

[55] S. Fazli, F. Popescu, M. Danóczy, B. Blankertz, K.-R. Müller, and C. Grozea, "Subject-independent mental state classification in single trials," *Neural Networks*, vol. 22, no. 9, pp. 1305–1312, 2009.

[56] A. Jackson, C. T. Moritz, J. Mavoori, T. H. Lucas, and E. E. Fetz, "The neurochip BCI: towards a neural prosthesis for upper limb function," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 187–190, 2006.

[57] N. Birbaumer, "Breaking the silence: Brain-computer interfaces (BCI) for communication and motor control," *Psychophysiology*, vol. 43, pp. 517–532, 2006.

[58] J. Wolpaw and E. Wolpaw, *Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, 2012.

[59] F. Cincotti, D. Mattia, F. Aloise, S. Bufalari, G. Schalk, G. Oriolo, A. Cherubini, M. G. Marciani, and F. Babiloni, "Non-invasive brain-computer interface system: Towards its application as assistive technology," *Brain Research Bulletin*, vol. 75, no. 6, pp. 796–803, 2008.

[60] E. Niedermeyer and F. L. da Silva, *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, 2005.

[61] G. Gargiulo, R. A. Calvo, P. Bifulco, M. Cesarelli, C. Jin, A. Mohamed, and A. van Schaik, "A new EEG recording system for passive dry electrodes," *Clinical Neurophysiology*, vol. 121, no. 5, pp. 686–693, 2010.

[62] S. Dähne, F. C. Meinecke, S. Haufe, J. Höhne, M. Tangermann, K.-R. Müller, and V. V. Nikulin, "Spoc: A novel framework for relating

the amplitude of neuronal oscillations to behaviorally relevant parameters," *NeuroImage*, vol. 86, pp. 111–122, 2014.

[63] S. Dähne, V. V. Nikulin, D. Ramírez, P. J. Schreier, K.-R. Müller, and S. Haufe, "Finding brain oscillations with power dependencies in neuroimaging data," *NeuroImage*, vol. 96, pp. 334–348, 2014.

[64] V. V. Nikulin, G. Nolte, and G. Curio, "Cross-frequency decomposition: A novel technique for studying interactions between neuronal oscillations with different frequencies," *Clinical Neurophysiology*, vol. 123, no. 7, pp. 1353–1360, 2012.

[65] D. L. Schomer and F. L. Da Silva, *Niedermeyer's electroencephalography: basic principles, clinical applications, and related fields.* Lippincott Williams & Wilkins, 2012.

[66] A. Bollimunta, Y. Chen, C. E. Schroeder, and M. Ding, "Neuronal mechanisms of cortical alpha oscillations in awake-behaving macaques," *The Journal of Neuroscience*, vol. 28, no. 40, pp. 9976–9988, 2008.

[67] G. J. M. van Boxtel, A. J. M. Denissen, M. Jäger, D. Vernon, M. K. J. Dekker, V. Mihajlović, and M. M. Sitskoorn, "A novel self-guided approach to alpha activity training," *International Journal of Psychophysiology*, vol. 83, no. 3, pp. 282–294, 2012.

[68] N. R. Cooper, R. J. Croft, S. J. J. Dominey, A. P. Burgess, and J. H. Gruzelier, "Paradox lost? exploring the role of alpha oscillations during externally vs. internally directed attention and the implications for idling and inhibition hypotheses," *International Journal of Psychophysiology*, vol. 47, no. 1, pp. 65–74, 2003.

[69] G. Pfurtscheller, C. Neuper, and W. Mohl, "Event-related desynchronization (ERD) during visual processing," *International Journal of Psychophysiology*, vol. 16, no. 2, pp. 147–153, 1994.

[70] W. Klimesch, M. Doppelmayr, T. Pachinger, and H. Russegger, "Event-related desynchronization in the alpha band and the processing of semantic information," *Cognitive Brain Research*, vol. 6, no. 2, pp. 83–94, 1997.

[71] W. Klimesch, "EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis," *Brain Research Reviews*, vol. 29, no. 2, pp. 169–195, 1999.

[72] G. Pfurtscheller, "Event-related synchronization (ERS): an electrophysiological correlate of cortical areas at rest," *Electroencephalography and Clinical Neurophysiology*, vol. 83, no. 1, pp. 62–69, 1992.

[73] P. Ritter, M. Moosmann, and A. Villringer, "Rolandic alpha and beta EEG rhythms' strengths are inversely related to fMRI-BOLD signal in primary somatosensory and motor cortex," *Human Brain Mapping*, vol. 30, no. 4, pp. 1168–1187, 2009.

[74] D. Huang and et al., "Decoding human motor activity from EEG single trials for a discrete two-dimensional cursor control," *Journal of Neural Engineering*, vol. 6, no. 4, p. 046005, 2009.

[75] C. Neuper, R. Scherer, M. Reiner, and G. Pfurtscheller, "Imagery of motor actions: Differential effects of kinesthetic and visual-motor mode of imagery in single-trial EEG," *Cognitive Brain Research*, vol. 25, no. 3, pp. 668–677, 2005.

[76] C. Neuper, R. Scherer, S. Wriessnegger, and G. Pfurtscheller, "Motor imagery and action observation: Modulation of sensorimotor brain

rhythms during mental control of a brain-computer interface," *Clinical Neurophysiology*, vol. 120, no. 2, pp. 239–247, 2009.

[77] J. A. Pineda, B. Z. Allison, and A. Vankov, "The effects of self-movement, observation, and imagination on μ rhythms and readiness potentials (RP's): toward a brain-computer interface (BCI)," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 219–222, 2000.

[78] G. Pfurtscheller and F. H. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles," *Clinical Neurophysiology*, vol. 110, no. 11, pp. 1842–1857, 1999.

[79] G. Bin, X. Gao, Z. Yan, B. Hong, and S. Gao, "An online multi-channel ssvep-based brain-computer interface using a canonical correlation analysis method," *Journal of Neural Engineering*, vol. 6, no. 4, pp. 1–6, 2009.

[80] M. Kaper, P. Meinicke, U. Grossekathoefer, T. Lingner, and H. Ritter, "BCI competition 2003-data set IIb: support vector machines for the P300 speller paradigm," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1073–1076, 2004.

[81] C. Y. Pang and M. M. Mueller, "Test-retest reliability of concurrently recorded steady-state and somatosensory evoked potentials in somatosensory sustained spatial attention," *Biological Psychology*, 2014.

[82] N. J. Hill, A. Moinuddin, S. Kienzle, A.-K. Häuser, and G. Schalk, "Communication and control by listening: towards optimal design of a two-class auditory streaming brain-computer interface," *Frontiers in Neuroscience*, vol. 6, 2012.

[83] C. Breitwieser, C. Pokorny, C. Neuper, and G. R. Muller-Putz, "Somatosensory evoked potentials elicited by stimulating two fingers from one hand-usable for BCI?" in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society,EMBC, 2011*, pp. 6373–6376.

[84] C. Zickler, A. Riccio, F. Leotta, S. Hillian-Tress, S. Halder, E. Holz, P. Staiger-Sälzer, E.-J. Hoogerwerf, L. Desideri, and D. Mattia, "A brain-computer interface as input channel for a standard assistive technology software," *Clinical EEG and Neuroscience*, vol. 42, no. 4, pp. 236–244, 2011.

[85] B. Blankertz, G. Dornhege, M. Krauledat, K. R. Müller, V. Kunzmann, F. Losch, and G. Curio, "The berlin brain-computer interface: EEG-based communication without subject training," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 147–152, 2006.

[86] F. Nijboer, A. Furdea, I. Gunst, J. Mellinger, D. J. McFarland, N. Birbaumer, and A. Kübler, "An auditory brain-computer interface (BCI)," *Journal of Neuroscience Methods*, vol. 167, no. 1, pp. 43–50, 2008.

[87] T. Solis-Escalante, G. Muller-Putz, C. Brunner, V. Kaiser, and G. Pfurtscheller, "Analysis of sensorimotor rhythms for the implementation of a brain switch for healthy subjects," *Biomedical Signal Processing and Control*, vol. 5, no. 1, pp. 15–20, 2010.

[88] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K. R. Muller, "Optimizing spatial filters for robust EEG single-trial analysis," *IEEE Signal Processing Magazine*, vol. 25, no. 1, pp. 41–56, 2008.

[89] A. Kubler, N. Neumann, J. Kaiser, B. Kotchoubey, T. Hinterberger, and N. P. Birbaumer, "Brain-computer communication: Self-regulation of slow cortical potentials for verbal communication," *Archives of Physical Medicine and Rehabilitation*, vol. 82, no. 11, pp. 1533–1539, 2001.

[90] J. Takahashi, A. Yasumura, E. Nakagawa, and M. Inagaki, "Changes in negative and positive EEG shifts during slow cortical potential training in children with attention-deficit/hyperactivity disorder: a preliminary investigation," *Neuroreport*, vol. 25, no. 8, pp. 618–624, 2014.

[91] B. Kotchoubey, S. Haisst, I. Daum, M. Schugens, and N. Birbaumer, "Learning and selfregulation of slow cortical potentials in older adults," *Experimental Aging Research*, vol. 26, no. 1, pp. 15–35, 2000.

[92] I. Daum, B. Rockstroh, N. Birbaumer, T. Elbert, A. Canavan, and W. Lutzenberger, "Behavioural treatment of slow cortical potentials in intractable epilepsy: neuropsychological predictors of outcome," *J Neurol Neurosurg Psychiatry*, vol. 56, no. 1, pp. 94–97, 1993.

[93] G. Pfurtscheller, C. Brunner, A. Schlögl, and F. H. Lopes da Silva, "Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks," *NeuroImage*, vol. 31, no. 1, pp. 153–159, 2006.

[94] G. Pfurtscheller, C. Neuper, D. Flotzinger, and M. Pregenzer, "EEG-based discrimination between imagination of right and left hand movement," *Electroencephalography and Clinical Neurophysiology*, vol. 103, no. 6, pp. 642–651, 1997.

[95] J. R. Wolpaw, D. J. McFarland, and T. M. Vaughan, "Brain-computer interface research at the wadsworth center," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 222–226, 2000.

[96] N. Sharma, V. M. Pomeroy, and J.-C. Baron, "Motor imagery: A backdoor to the motor system after stroke?" *Stroke*, vol. 37, no. 7, pp. 1941–1952, Jul. 2006.

[97] A. Dunsky, R. Dickstein, E. Marcovitz, S. Levy, and J. Deutsch, "Home-based motor imagery training for gait rehabilitation of people with chronic poststroke hemiparesis," *Archives of Physical Medicine and Rehabilitation*, vol. 89, no. 8, pp. 1580–1588, 2008.

[98] Y. Huijuan, G. Cuntai, C. Karen Sui Geok, C. See San, W. Chuan Chu, S. Phua Kok, T. Christina Ka Yin, and A. Kai Keng, "Detection of motor imagery of swallow EEG signals based on the dual-tree complex wavelet transform and adaptive model selection," *Journal of Neural Engineering*, vol. 11, no. 3, p. 035016, 2014.

[99] C. Kranczioch, C. Zich, I. Schierholz, and A. Sterr, "Mobile EEG and its potential to promote the theory and application of imagery-based motor rehabilitation," *International Journal of Psychophysiology*, vol. 91, no. 1, pp. 10–15, 2014.

[100] S. Makeig, A. J. Bell, T.-P. Jung, and T. J. Sejnowski, "Independent component analysis of electroencephalographic data," *Advances in neural information processing systems*, pp. 145–151, 1996.

[101] W. Zhou and J. Gotman, "Automatic removal of eye movement artifacts from the EEG using ICA and the dipole model," *Progress in Natural Science*, vol. 19, no. 9, pp. 1165–1170, 2009.

[102] D. J. McFarland, L. M. McCane, S. V. David, and J. R. Wolpaw, "Spatial filter selection for EEG-based communication," *Electroencephalography and Clinical Neurophysiology*, vol. 103, no. 3, pp. 386–394, 1997.

[103] K. Fukunaga, *Introduction to statistical pattern recognition.* Academic press, 1990.

[104] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Transactions on Rehabilitation Engineering,*, vol. 8, no. 4, pp. 441–446, 2000.

[105] W. Samek, F. C. Meinecke, and K. R. Muller, "Transferring subspaces between subjects in brain-computer interfacing," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 8, pp. 2289–2298, 2013.

[106] M. Arvaneh, C. Guan, K. K. Ang, and C. Quek, "EEG data space adaptation to reduce intersession nonstationarity in brain-computer interface," *Neural Computation*, vol. 25, no. 8, pp. 2146–2171, 2013.

[107] D. J. McFarland, C. W. Anderson, K. R. Muller, A. Schlogl, and D. J. Krusienski, "BCI meeting 2005-workshop on BCI signal processing: feature extraction and translation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 135–138, 2006.

[108] A. Schloegl, K. Lugger, and G. Pfurtscheller, "Using adaptive autoregressive parameters for a brain-computer-interface experiment," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 4, 1997, pp. 1533–1535 vol.4.

[109] G. Pfurtscheller and C. Neuper, "Motor imagery and direct brain-computer communication," in *Proceedings of the IEEE*, vol. 89, 2001, pp. 1123–1134.

[110] D. Coyle, G. Prasad, and T. McGinnity, "A time-frequency approach to feature extraction for a brain-computer interface with a comparative

analysis of performance measures," *EURASIP Journal on Advances in Signal Processing*, vol. 2005, no. 19, p. 861614, 2005.

[111] V. J. Samar, A. Bopardikar, R. Rao, and K. Swartz, "Wavelet analysis of neuroelectric waveforms: A conceptual tutorial," *Brain and Language*, vol. 66, no. 1, pp. 7–60, 1999.

[112] F. Lotte, M. Congedo, A. LÂŽecuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for EEG-based braincomputer interfaces," *Neural Engineering*, vol. 4, no. 2, pp. R1–R13, 2007.

[113] Y. Li and C. Guan, "Joint feature re-extraction and classification using an iterative semi-supervised support vector machine algorithm," *Machine Learning*, vol. 71, no. 1, pp. 33–53, Apr. 2008.

[114] J. Yingying, W. Xiaopei, and G. Xiaojing, "Motor imagery classification based on the optimized SVM and BPNN by GA," in *International Conference on Intelligent Control and Information Processing (ICICIP)*, 2010, pp. 344–347.

[115] S. Hongyu, X. Yang, S. Yaoru, Z. Huaping, and Z. Jinhua, "Online EEG classification for brain-computer interface based on CSP and SVM," in *International Congress on Image and Signal Processing (CISP)*, vol. 9, 2010, pp. 4105–4108.

[116] A. Bamdadian, G. Cuntai, A. Kai Keng, and X. Jianxin, "Real coded GA-based SVM for motor imagery classification in a brain-computer interface," in *9th IEEE International Conference on Control and Automation (ICCA), 2011*, pp. 1355–1359.

[117] K. K. Ang, Z. Y. Chin, C. Wang, C. Guan, and H. Zhang, "Filter Bank Common Spatial Pattern algorithm on BCI Competition IV Datasets 2a and 2b," *Frontiers in Neuroscience*, vol. 6, 2012.

[118] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, "Filter bank common spatial pattern (FBCSP) in brain-computer interface," in *IEEE International Joint Conference on Neural Networks*, pp. 2390–2397.

[119] D. J. McFarland, L. M. McCane, and J. R. Wolpaw, "EEG-based communication and control: short-term role of feedback," *IEEE Transactions on Rehabilitation Engineering*, vol. 6, no. 1, pp. 7–11, 1998.

[120] M. Gonzalez-Franco, Y. Peng, Z. Dan, H. Bo, and G. Shangkai, "Motor imagery based brain-computer interface: A study of the effect of positive and negative feedback," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 6323–6326.

[121] L. Kalra, "Stroke rehabilitation 2009: Old chestnuts and new insights," *Stroke*, vol. 41, pp. e88–e90, Feb. 2010.

[122] J. J. Daly, R. Cheng, K. H. J. M. Rogers, and K. L. M. E. Dohring, "Development and testing of non-invasive BCI+FES/robot system for use in motor re-learning after stroke," in *proceeding of 13th Annual Conference of the International Functional Electrical Stimulation Society "From Movement to Mind"*, pp. 200–202.

[123] O. Lambercy, L. Dovat, R. Gassert, E. Burdet, T. Chee Leong, and T. Milner, "A haptic knob for rehabilitation of hand function," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 15, no. 3, pp. 356–366, 2007.

[124] D. Intiso, V. Santilli, M. Grasso, R. Rossi, and I. Caruso, "Rehabilitation of walking with electromyographic biofeedback in foot-drop after stroke," *Stroke*, vol. 25, pp. 1189–1192, Jun. 1994.

[125] A. a. b. Dunsky, R. a. Dickstein, C. c. Ariav, J. d. Deutsch, and E. c. Marcovitz, "Motor imagery practice in gait rehabilitation of chronic

post-stroke hemiparesis: four case studies," *International Journal of Rehabilitation Research*, vol. 29, no. 4, pp. 351–356, 2006, report.

[126] J. E. Deutsch and R. Dickstein, "Motor imagery in physical therapist practice," *Physical Therapy*, vol. 87, no. 7, pp. 942–953, 2007.

[127] J. R. Wolpaw and D. J. McFarland, "Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 51, pp. 17 849–17 854, 2004.

[128] R. Leeb, D. Friedman, G. R. Müller-Putz, R. Scherer, M. Slater, and G. Pfurtscheller, "Self-paced (asynchronous) BCI control of a wheelchair in virtual environments: a case study with a tetraplegic," *Computational Intelligence and Neuroscience*, vol. 2007, 2007.

[129] G. Pfurtscheller, C. Guger, G. Müller, G. Krausz, and C. Neuper, "Brain oscillations control hand orthosis in a tetraplegic," *Neuroscience Letters*, vol. 292, no. 3, pp. 211–214, 2000.

[130] A. J. Doud, J. P. Lucas, M. T. Pisansky, and B. He, "Continuous three-dimensional control of a virtual helicopter using a motor imagery based brain-computer interface," *PLoS ONE*, vol. 6, no. 10, p. e26322, 2011.

[131] A. S. Royer, A. J. Doud, M. L. Rose, and H. Bin, "EEG control of a virtual helicopter in 3-dimensional space using intelligent control strategies," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 6, pp. 581–589, 2010.

[132] L. Karl, C. Kaitlin, D. Alexander, S. Kaleb, R. Eitan, and H. Bin, "Quadcopter control in three-dimensional space using a noninvasive motor imagery-based brain-computer interface," *Journal of Neural Engineering*, vol. 10, no. 4, p. 046003, 2013.

[133] C. Vidaurre, C. Sannelli, K.-R. Müller, and B. Blankertz, *Machine-Learning Based Co-adaptive Calibration: A Perspective to Fight BCI Illiteracy*, ser. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2010, vol. 6076, pp. 413–420.

[134] C. Vidaurre and B. Blankertz, "Towards a cure for BCI illiteracy," *Brain Topography*, vol. 23, no. 2, pp. 194–198, 2010.

[135] N. Neumann and N. Birbaumer, "Predicors of successful self control during brain computer communication," *Neurology, Neurosurgery and Psychiatry*, vol. 74, pp. 1117–1121, 2003.

[136] A. kübler, N. Neumann, B. Wilhelm, T. Hinterberger, and N. Birbaumer, "Predictability of brain-computer communication," *Psychophysiology*, vol. 18, pp. 121–129, 2004.

[137] M. Grosse-Wentrup, B. Scholkopf, and J. Hill, "Causal influence of gamma oscillations on the sensorimotor rhythm," *NeuroImage*, vol. 56, pp. 837–842, 2011.

[138] E. M. Hammer, S. Halder, B. Blankertz, C. Sannelli, T. Dickhaus, S. Kleih, K.-R. Müller, and A. Kübler, "Psychological predictors of SMR-BCI performance," *Biological Psychology*, vol. 89, no. 1, pp. 80–86, 2012.

[139] C. L. Maeder, C. Sannelli, S. Haufe, and B. Blankertz, "Pre-stimulus sensorimotor rhythms influence brain-computer interface classification performance," *IEEE Transactions on Neural Systems and Rehabilitation Engineering,*, vol. 20, no. 5, pp. 653–662, 2012.

[140] H. van Dijk, J.-M. Schoffelen, R. Oostenveld, and O. Jensen, "Prestimulus oscillatory activity in the alpha band predicts visual discrimi-

nation ability," *The Journal of Neuroscience*, vol. 28, no. 8, pp. 1816–1823, 2008.

[141] J. Fell, E. Ludowig, B. P. Staresina, T. Wagner, T. Kranz, C. E. Elger, and N. Axmacher, "Medial temporal Theta/Alpha power enhancement precedes successful memory encoding: Evidence based on intracranial EEG," *The Journal of Neuroscience*, vol. 31, no. 14, pp. 5392–5397, 2011.

[142] A. R. Haig and E. Gordon, "Prestimulus EEG alpha phase synchronicity influences N100 amplitude and reaction time," *Psychophysiology*, vol. 35, no. 5, pp. 591–595, 1998.

[143] C. Tangwiriyasakul, R. Verhagen, M. J. A. M. v. Putten, and W. L. C. Rutten, "Importance of baseline in event-related desynchronization during a combination task of motor imagery and motor observation," *Journal of Neural Engineering*, vol. 10, no. 2, p. 026009, 2013.

[144] V. Romei, J. Gross, and G. Thut, "On the role of prestimulus alpha rhythms over occipito-parietal areas in visual input regulation: Correlation or causation?" *The Journal of Neuroscience*, vol. 30, no. 25, pp. 8692–8697, 2010.

[145] R. A. Rensink, J. K. O'Regan, and J. J. Clark, "To see or not to see: The need for attention to perceive changes in scenes," *Psychological Science*, vol. 8, no. 5, pp. 368–373, 1997.

[146] S. Guderian, B. H. Schott, A. Richardson-Klavehn, and E. Düzel, "Medial temporal theta state before an event predicts episodic encoding success in humans," *Proceedings of the National Academy of Sciences*, vol. 106, no. 13, pp. 5365–5370, 2009.

[147] P. Missonnier, M. P. Deiber, G. Gold, P. Millet, M. Gex-Fabry Pun, L. Fazio-Costa, P. Giannakopoulos, and V. Ibáñez, "Frontal theta event-related synchronization: comparison of directed attention and working memory load effects," *Journal of Neural Transmission*, vol. 113, no. 10, pp. 1477–1486, 2006.

[148] S. A. A. Massar, V. Rossi, D. J. L. G. Schutter, and J. L. Kenemans, "Baseline EEG theta/beta ratio and punishment sensitivity as biomarkers for feedback-related negativity (FRN) and risk-taking," *Clinical Neurophysiology*, vol. 123, no. 10, 2012.

[149] R. J. Barry, A. R. Clarke, S. J. Johnstone, R. McCarthy, and M. Selikowitz, "Electroencephalogram θ/β ratio and arousal in attention-deficit/hyperactivity disorder: Evidence of independent processes," *Biological Psychiatry*, vol. 66, no. 4, pp. 398–401, 2009.

[150] B. Lou, Y. Li, M. G. Philiastides, and P. Sajda, "Prestimulus alpha power predicts fidelity of sensory encoding in perceptual decision making," *NeuroImage*, vol. 87, pp. 242–251, 2014.

[151] V. Wyart and C. Tallon-Baudry, "How ongoing fluctuations in human visual cortex predict perceptual awareness: Baseline shift versus decision bias," *The Journal of Neuroscience*, vol. 29, no. 27, pp. 8715–8725, 2009.

[152] P. Sauseng, W. Klimesch, W. Stadler, M. Schabus, M. Doppelmayr, S. Hanslmayr, W. R. Gruber, and N. Birbaumer, "A shift of visual spatial attention is selectively associated with human EEG alpha activity," *European Journal of Neuroscience*, vol. 22, no. 11, pp. 2917–2926, 2005.

[153] B. Blankertz, F. Losch, M. Krauledat, G. Dornhege, G. Curio, and K. R. Müller, "The berlin brain-computer interface: Accurate perfor-

mance from first-session in BCI-naive subjects," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 10, pp. 2452–2462, 2008.

[154] V. V. Nikulin, G. Nolte, and G. Curio, "A novel method for reliable and fast extraction of neuronal EEG/MEG oscillations on the basis of spatio-spectral decomposition," *NeuroImage*, vol. 55, no. 4, pp. 1528–1535, 2011.

[155] K. Fukunaga, *Introduction to Statistical Pattern Recognition*. Elsevier Science, 1990.

[156] L. C. Parra, C. D. Spence, A. D. Gerson, and P. Sajda, "Recipes for the linear analysis of EEG," *Neuroimage*, vol. 28, no. 2, pp. 326–341, 2005.

[157] J. H. Gruzelier, "EEG-neurofeedback for optimising performance. I: A review of cognitive and affective outcome in healthy participants," *Neuroscience & Biobehavioral Reviews*, 2014.

[158] ——, "EEG-neurofeedback for optimising performance. II: Creativity, the performing arts and ecological validity," *Neuroscience & Biobehavioral Reviews*, 2014.

[159] ——, "EEG-neurofeedback for optimising performance. III: A review of methodological and theoretical considerations," *Neuroscience & Biobehavioral Reviews*, 2014.

[160] N. Duric, J. Assmus, D. Gundersen, and I. Elgen, "Neurofeedback for the treatment of children and adolescents with adhd: a randomized and controlled clinical trial using parental reports," *BMC Psychiatry*, vol. 12, no. 1, p. 107, 2012.

[161] S. E. Kober, M. Witte, M. Ninaus, C. Neuper, and G. Wood, "Learning to modulate one's own brain activity: The effect of spontaneous mental strategies," *Frontiers in Human Neuroscience*, vol. 7, 2013.

[162] K. Hoedlmoser, T. Pecherstorfer, G. Gruber, P. Anderer, M. Doppelmayr, W. Klimesch, and M. Schabus, "Instrumental conditioning of human sensorimotor rhythm (12-15 Hz) and its impact on sleep as well as declarative learning," *Sleep*, vol. 31, no. 10, pp. 1401–8, 2008.

[163] N. Wenya, W. Feng, I. V. Mang, and R. Agostinho, "Resting alpha activity predicts learning ability in alpha neurofeedback," *Frontiers in Human Neuroscience*, 2014.

[164] M. Witte, S. E. Kober, M. Ninaus, C. Neuper, and G. Wood, "Control beliefs can predict the ability to up-regulate sensorimotor rhythm during neurofeedback training," *Frontiers in Human Neuroscience*, vol. 7, 2013.

[165] G. Pfurtscheller, C. Neuper, G. R. Muller, B. Obermaier, G. Krausz, A. Schlogl, R. Scherer, B. Graimann, C. Keinrath, D. Skliris, M. Wortz, G. Supp, and C. Schrank, "Graz-BCI: state of the art and clinical applications," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 11, no. 2, pp. 1–4, 2003.

[166] G. R. Müller-Putz, V. Kaiser, T. Solis-Escalante, and G. Pfurtscheller, "Fast set-up asynchronous brain-switch based on detection of foot motor imagery in 1 -channel EEG," *Medical & Biological Engineering & Computing*, vol. 48, pp. 229–233, 2010.

[167] M. Alegre, A. Labarga, I. G. Gurtubay, J. Iriarte, A. Malanda, and J. Artieda, "Beta electroencephalograph changes during passive movements: sensory afferences contribute to beta event-related desynchro-

nization in humans," *Neuroscience Letters*, vol. 331, no. 1, pp. 29–32, 2002.

[168] G. R. Müller, C. Neuper, R. Rupp, C. Keinrath, H. J. Gerner, and G. Pfurtscheller, "Event-related beta EEG changes during wrist movements induced by functional electrical stimulation of forearm muscles in man," *Neuroscience Letters*, vol. 340, no. 2, pp. 143–147, 2003.

[169] K. K. Ang, C. Guan, C. Wang, K. S. Phua, T. A. H. G., and Z. Y. Chin, "Calibrating EEG-based motor imagery brain-computer interface from passive movement," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 4199–4202.

[170] M. Sugiyama, T. Suzuki, S. Nakajima, H. Kashima, P. von Bünau, and M. Kawanabe, "Direct importance estimation for covariate shift adaptation," *Annals of the Institute of Statistical Mathematics*, vol. 60, no. 4, pp. 699–746, 2008.

[171] M. Sugiyama, M. Krauledat, and K.-R. Müller, "Covariate shift adaptation by importance weighted cross validation," *Machine Learning Research*, vol. 8, pp. 985–1005, 2007.

[172] Z. Qibin, Z. Liqing, A. Cichocki, and L. Jie, "Incremental common spatial pattern algorithm for BCI," in *IEEE International Joint Conference on Neural Networks (IJCNN)*, 2008, pp. 2656–2659.

[173] W. Samek, C. Vidaurre, K.-R. Müller, and M. Kawanabe, "Stationary common spatial patterns for brain-computer interfacing," *Journal of Neural Engineering*, vol. 9, no. 2, p. 026013, 2012.

[174] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, "Mutual information-based selection of optimal spatial and temporal patterns for single- trial

EEG-based BCIs," *Pattern Recognition*, vol. 45, no. 6, pp. 2137–2144, 2012.

[175] K. Hyohyeong, N. Yunjun, and C. Seungjin, "Composite common spatial pattern for subject-to-subject transfer," *IEEE Signal Processing Letters*, vol. 16, no. 8, pp. 683–686, 2009.

[176] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*. John Wiley and Sons Inc, 2001.

[177] I. Cohen, F. G. Cozman, N. Sebe, M. C. Cirelo, and T. S. Huang, "Semisupervised learning of classifiers: theory, algorithms, and their application to human-computer interaction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 12, pp. 1553–1566, 2004.

[178] A. Schlogl and C. Brunner, "BioSig: a free and open source software library for BCI research," *Computer*, vol. 41, no. 10, pp. 44–50, 2008.

[179] A. Y. Kaplan, A. A. Fingelkurts, A. A. Fingelkurts, S. V. Borisov, and B. S. Darkhovsky, "Nonstationary nature of the brain activity as revealed by EEG/MEG methodological, practical and conceptual challenges," *Signal Processing*, vol. 85, no. 11, pp. 2190–2212, 2005.

[180] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, no. 1, pp. 489–501, 2006.

[181] X. Liu, C. Gao, and P. Li, "A comparative analysis of support vector machines and extreme learning machines," *Neural Networks*, vol. 33, pp. 58–66, 2012.

[182] Y. Song, J. Crowcroft, and J. Zhang, "Automatic epileptic seizure detection in EEGs based on optimized sample entropy and extreme learning machine," *Journal of neuroscience methods*, vol. 210, no. 2, pp. 132–146, 2012.

[183] R. Minhas, A. A. Mohammed, and Q. M. Wu, "Incremental learning in human action recognition based on snippets," *IEEE Transactions on Circuits and Systems for Video Technology,*, vol. PP, no. 99, pp. 1–1, 2011.

[184] K. Choi, K.-A. Toh, and H. Byun, "Incremental face recognition for large-scale social network services," *Pattern Recognition*, vol. 45, no. 8, pp. 2868–2883, 2012.

[185] L. Yan, H. Kambara, Y. Koike, and M. Sugiyama, "Application of covariate shift adaptation techniques in brain computer interfaces," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 6, pp. 1318–1324, 2010.

[186] G. Ferber, "Treatment of some nonstationarities in the EEG," *Neuropsychobiology*, vol. 17, pp. 100–104, 1987.

[187] C. Vidaurre, A. Schlogl, R. Cabeza, R. Scherer, and G. Pfurtscheller, "A fully on-line adaptive BCI," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 6, pp. 1214–1219, 2006.

[188] L. Sidath Ravindra, G. Cuntai, Z. Haihong, A. Kai Keng, X. JianXin, and L. Tong Heng, "Dynamically weighted ensemble classification for non-stationary EEG processing," *Journal of Neural Engineering*, vol. 10, no. 3, p. 036007, 2013.

Publications and Awards

The contents of this thesis are based on the following papers that have been published, accepted, or submitted to the peer-reviewed journals and conferences.

Journal paper

[1] A. Bamdadian, C. Guan, K. K. Ang, J. Xu, The predictive role of pre- cue EEG rhythms on MI-based BCI classification performance, Journal of Neuroscience Methods 235 (2014) 138-144.

Conference papers

[1] Atieh Bamdadian, Cuntai Guan, Kai Keng Ang, Jianxin Xu, Towards improvement of MI-BCI performance of subjects with BCI deficiency, International IEEE/EMBS Conference on Neural Engineering (NER), 2015

[2] Atieh Bamdadian, Cuntai Guan, Kai Keng Ang, Jianxin Xu, Predicting the performance of the MI-based BCI based on pre-cue EEG data, 5th BCI Meeting, 2013

[3] Atieh Bamdadian, Cuntai Guan, Kai Keng Ang, Jianxin Xu, Improving session-to-session transfer performance of motor imagery-based BCI using Adaptive Extreme Learning Machine, International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2013

[4] Atieh Bamdadian, Cuntai Guan, Kai Keng Ang, Jianxin Xu, Online Semi-supervised Learning with KL Distance Weighting for Motor Imagery-based BCI, International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2012

[5] Atieh Bamdadian, Cuntai Guan, Kai Keng Ang, Jianxin Xu, Real Coded GA-Based SVM for Motor Imagery Classification in a Brain-Computer Interface, IEEE International Conference on Control and Automation (ICCA), 2011

International awards

1. The most innovative poster: Atieh Bamdadian, Cuntai Guan, Kai Keng Ang, Jianxin Xu, Predicting the performance of the MI-based BCI based on pre-cue EEG data. 5th BCI Meeting, 2013.