

CLASSIFICATION OF MOTOR IMAGERY TASKS IN EEG SIGNAL AND
ITS APPLICATION TO A BRAIN-COMPUTER INTERFACE FOR
CONTROLLING ASSISTIVE ENVIRONMENTAL DEVICES

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ERMAN ACAR

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

FEBRUARY 2011

Approval of the thesis

**CLASSIFICATION OF MOTOR IMAGERY TASKS IN EEG SIGNAL
AND ITS APPLICATION TO A BRAIN-COMPUTER INTERFACE FOR
CONTROLLING ASSISTIVE ENVIRONMENTAL DEVICES**

submitted by **ERMAN ACAR** in partial fulfillment of the requirements for the degree of **Master of Science in Electrical and Electronics Engineering, Middle East Technical University** by,

Prof. Dr. Canan Özgen
Dean, Graduate School of **Natural and Applied Sciences** _____

Prof. Dr. İsmet Erkmn
Head of Department, **Electrical and Electronics Engineering** _____

Prof. Dr. Nevzat Güneri Gençer
Supervisor, **Electrical and Electronics Engineering, METU** _____

Examining Committee Members:

Prof. Dr. Murat Eyüboğlu
Electrical and Electronics Engineering, METU _____

Prof. Dr. Nevzat Güneri Gençer
Electrical and Electronics Engineering, METU _____

Assist. Prof. Dr. Yeşim Serinağaoğlu Doğrusöz
Electrical and Electronics Engineering, METU _____

Prof. Dr. Uğur Halıcı
Electrical and Electronics Engineering, METU _____

Assoc. Prof. Dr. Süha Yağcıoğlu
Biophysics Dept., Faculty of Medicine, Hacettepe University _____

Date: 14.02.2011

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: Erman Acar

Signature :

ABSTRACT

CLASSIFICATION OF MOTOR IMAGERY TASKS IN EEG SIGNAL AND ITS APPLICATION TO A BRAIN- COMPUTER INTERFACE FOR CONTROLLING ASSISTIVE ENVIRONMENTAL DEVICES

Acar, Erman

M.Sc., Department of Electrical and Electronics Engineering

Supervisor: Prof. Dr. Nevzat Güneri Gençer

February 2011, 92 Pages

This study focuses on realization of a Brain Computer Interface (BCI) for the paralyzed to control assistive environmental devices. For this purpose, different motor imagery tasks are classified using different signal processing methods. Specifically, band-pass filtering, Laplacian filtering, and common average reference (CAR) filtering are used to enhance the EEG signal. For feature extraction; Common Spatial Pattern (CSP), Power Spectral Density (PSD), and Principal Component Analysis (PCA) are tested. Linear Feature Normalization (LFN), Gaussian Feature Normalization (GFN), and Unit-norm Feature Vector Normalization (UFVN) are studied in Support Vector Machine (SVM) and Artificial Neural Network (ANN) classification. In order to evaluate and compare the performance of the methodologies, classification accuracy, Cohen's kappa coefficient, and Nykopp's information transfer are utilized.

The first experiments on classifying motor imagery tasks are realized on the 3-class dataset (V) provided for BCI Competition III. Also, a 4-class problem is studied using the dataset (IIa) provided for BCI Competition IV. Then, 5 different tasks are studied in the METU Brain Research Laboratory to find the optimum number and type of tasks to control a motor imagery based BCI. Thereafter, an interface is designed for the paralyzed to control assistive environmental devices. Finally, a test application is implemented and online performance of the design is evaluated.

Keywords: Brain Computer Interface, BCI, Electroencephalography, EEG, Environmental Control, Motor Imagery, Event Related Desynchronization - Synchronization (ERD - ERS), Power Spectral Density, Common Spatial Patterns, Support Vector Machines, Artificial Neural Networks.

ÖZ

EEG SİNYALLERİNDEKİ HAREKET DÜŞÜNSEL GÖREVLERİN SINIFLANDIRILMASI VE YARDIMCI ÇEVRESEL CİHAZLARI KONTROL İÇİN BİR BEYİN BİLGİSAYAR ARAYÜZÜNE UYGULANMASI

Acar, Erman

Yüksek Lisans, Elektrik-Elektronik Mühendisliği Bölümü

Tez Yöneticisi: Prof. Dr. Nevzat Güneri Gençer

Şubat 2011, 92 Sayfa

Bu çalışma felçli hastaların yardımcı çevresel cihazları kontrolü için bir Beyin Bilgisayar Arayüzü gerçekleştirmeye odaklanmıştır. Bu amaçla, farklı hareket düşünsel görevler, farklı işaret işleme yöntemleri kullanılarak sınıflandırılmışlardır. Özellikle, EEG işaretini iyileştirmek için; bant geçiren süzgeç, Laplace süzgeç ve Genel Ortalama Referans (GOR) süzgeci kullanılmıştır. Öznitelik çıkartmak için, Ortak Uzamsal Örüntü (OUÖ), Spektral Güç Yoğunluğu (SGY), Ana Bileşenler Analizi (ABA) test edilmiştir. Destek Vektör Makinaları (DVM) ve Yapay Sinir Ağları (YSA) ile sınıflandırmada Doğrusal Öznitelik Düzgeleme (DÖD), Gauss Öznitelik Düzgeleme (GÖD) ve Birim-düzge Öznitelik Vektörü Düzgeleme (BÖVD) çalışılmıştır. Yöntemlerin başarımlarını ölçmek ve karşılaştırmak için sınıflandırma doğruluğundan, Cohen'in kappa katsayısından ve Nykopp'un bilgi aktarımından faydalanılmıştır.

Hareket dűşűnsel gűrevlerin sınıflandırılmasına yűnelik ilk deneyler 3. BBA Yarışması iin saėlanan 3-sınıflı veri kűmesi (V) ile gerekleřtirilmiřtir. Ayrıca, 4. BBA Yarışması iin saėlanan 4-sınıflı veri kűmesi (IIa) de alıřılmıřtır. Daha sonra hareket dűşűnsel bir BBA'yı kontrol etmek iin en iyi gűrev tűrű ve eřidini belirlemeye yűnelik, ODTŲ Beyin Arařtırmaları Laboratuvarında 5 farklı gűrev alıřılmıřtır. Sonra, felli hastaların yardımcı evresel cihazları kontrol etmesi iin bir arayűz tasarlanmıřtır. Son olarak, bu tasarımın evrimii bařarımını űlmeye yűnelik bir test uygulaması gerekleřtirilmiřtir.

Anahtar Sűzcűkler: Beyin Bilgisayar Arayűzű, BBA, Elektroensefalografi, EEG, evresel KontrolEnvironmental, Hareket Dűřűncesi, Olay İliřkili Desenkronizasyon - Senkronizasyon (OID – OIS), Spektral Gű Yoėunluėu, Ortak Uzamsal Őrűntű, Destek Vektűr Makinaları, Yapay Sinir Aėları.

To my Family

ACKNOWLEDGEMENTS

I would first like to thank my supervisor Prof. Dr. Nevzat Gençer for his guidance and support that have been invaluable for my career development.

I would like to acknowledge the debt I owe to my colleagues, Balkar Erdoğan and Berna Akıncı for their cooperation and support during the work.

Also, I would like to express my special thanks to Dr. Özlem Birgül, Dr. Ufuk Sakarya and Soner Büyükatalay for their advices and encouragement in my academic career.

I would like to thank to Assoc. Prof. Dr. Süha Yağcıođlu for sharing his laboratory facilities.

Finally, I would like to thank to TÜBİTAK-UZAY family for their friendship and support in any respect.

TABLE OF CONTENTS

ABSTRACT.....	iv
ÖZ	vi
ACKNOWLEDGEMENTS	ix
TABLE OF CONTENTS.....	x
LIST OF TABLES	xiii
LIST OF FIGURES	xv
CHAPTERS	
1 INTRODUCTION	1
2 BRAIN COMPUTER INTERFACES	4
2.1 <i>Signal Acquisition</i>	5
2.1.1 <i>Electroencephalography (EEG)</i>	7
2.1.2 <i>Magnetoencephalography (MEG)</i>	8
2.1.3 <i>Other Signal Acquisition Techniques</i>	9
2.2 <i>Neurophysiologic Signals</i>	10
2.2.1 <i>Evoked Responses</i>	10
2.2.1.1 <i>P300</i>	10
2.2.1.2 <i>Steady-State Visual Evoked Potential (SSVEP)</i>	11
2.2.1.3 <i>Slow Cortical Potentials (SCPs)</i>	12
2.2.2 <i>Induced Responses</i>	13

2.2.2.1	<i>Sensorimotor Rhythms (SMR)</i>	13
2.2.2.2	<i>Responses to Mental Tasks</i>	14
2.3	<i>Signal Processing</i>	14
2.4	<i>BCI Applications</i>	15
2.4.1	<i>Communication</i>	15
2.4.2	<i>Environmental Control</i>	16
2.4.3	<i>Movement Control</i>	17
2.4.4	<i>Locomotion</i>	18
2.4.5	<i>Neurorehabilitation</i>	19
3	SIGNAL PROCESSING IN BRAIN COMPUTER INTERFACES	20
3.1	<i>Signal Enhancement</i>	21
3.2	<i>Feature Extraction</i>	22
3.2.1	<i>Common Spatial Pattern (CSP)</i>	23
3.2.2	<i>Power Spectral Density (PSD)</i>	25
3.2.3	<i>Principal Component Analysis (PCA)</i>	26
3.3	<i>Normalization</i>	28
3.4	<i>Classification</i>	29
3.4.1	<i>Support Vector Machines (SVMs)</i>	29
3.4.2	<i>Artificial Neural Networks (ANNs)</i>	34
3.4.2.1	<i>Feed-forward Operation</i>	35
3.4.2.2	<i>Learning</i>	36
3.5	<i>Evaluation</i>	37
3.5.1	<i>The Confusion Matrix</i>	37
3.5.2	<i>Classification Accuracy</i>	38
3.5.3	<i>Cohen's Kappa Coefficient</i>	38
3.5.4	<i>Nykopp's Information Transfer</i>	39
4	EXPERIMENTS AND RESULTS	41

4.1 Experiments on BCI Competition III: Dataset V.....	41
4.1.1 Explanation of the Experiment	41
4.1.2 Precomputed Features.....	43
4.1.2.1 Explanation of the features	43
4.1.2.2 Results on Precomputed Features	43
4.1.3 Raw EEG Signals	54
4.1.3.1 Explanation of the features	54
4.1.3.2 Results on Raw EEG signals.....	54
4.1.4 Conclusion.....	58
4.2 Experiments on BCI Competition IV: Dataset IIa.....	59
4.2.1 Explanation of the Experiment	59
4.2.2 Explanation of the Data.....	61
4.2.2.1 Results.....	62
4.2.3 Conclusion.....	65
4.3 METU Brain Research Laboratory BCI Experiments.....	66
4.3.1 Offline Experiments	66
4.3.1.1 Explanation of the Experiment.....	67
4.3.1.2 Evaluation of the Data.....	68
4.3.1.3 Results.....	69
4.3.2 Online Experiments	77
5 CONCLUSION.....	83
REFERENCES	86

LIST OF TABLES

TABLES

Table 4-2 : Validation accuracies for all subjects and different γ & C combinations. The PCA-Coefficient is 99 and unit norm feature vector normalization and no feature normalization is used.	45
Table 4-3 : The validation accuracies for different normalization methods and γ &C combinations for subject 1 with PCA coefficient 99.....	46
Table 4-4 : The validation accuracies for different PCA-Coefficients and γ &C combinations for subject 1. The feature normalization type is Gaussian and no feature vector normalization is employed.....	47
Table 4-5 : Feature vector sizes for different PCA-Coefficients.	48
Table 4-6 : The validation accuracies for different subjects and nHidden & PCA-Coefficient combinations. Linear feature normalization and unit norm feature vector normalization is used.	49
Table 4-7 : The number of parameters and total number of combinations.....	52
Table 4-8 : Classification accuracy calculated using the responses computed 16 times per second.....	53
Table 4-9 : Classification accuracy calculated using the responses computed 2 times per second from the average of 8 consecutive outputs.....	53
Table 4-10 : The signal processing techniques used to classify the CSP features extracted from the raw EEG signal.....	55
Table 4-11 : Evaluation results with SVM for different filtering and normalization techniques	56
Table 4-12 : Evaluation results with ANN for different filtering and normalization techniques	57
Table 4-13 : The results of BCI competition III	58

Table 4-14 : Validation kappa values for subject A01 with different γ - C - p combinations. (spatial filtering: CAR, temporal filtering: 6-8Hz band-pass, feature normalization: linear, feature vector normalization: No)	63
Table 4-15 : The signal processing techniques used to classify the CSP features extracted from the BCI Competition IV data.....	64
Table 4-16 : The kappa coefficient values for different subject, filtering, and normalization types.....	65
Table 4-17 : The results of BCI competition IV	66
Table 4-18 : Validation accuracy results obtained for each validation process with different runs.....	70
Table 4-19 : Validation kappa coefficient results obtained for each validation process with different runs.....	71
Table 4-20: Validation information transfer results obtained for each validation process with different runs.....	72
Table 4-21 : Average (5-fold cross validation) classification accuracies for different motor imagery task combinations	73
Table 4-22 : Average (5-fold cross validation) kappa values for different motor imagery task combinations	74
Table 4-23 : Average (5-fold cross validation) information transfer values for different motor imagery task combinations	75
Table 4-24 : The elements and commands for the menu designed.....	77
Table 4-25 : The accuracy results of the online experiments.	82

LIST OF FIGURES

FIGURES

Figure 2-1 : The brain computer interface cycle [6].	5
Figure 2-2 : The spatial and temporal resolution scale of the signal acquisition techniques used in the BCI study (EEG: electroencephalography, MEG: magnetoencephalography, NIRS: near-infrared spectroscopy, fMRI: functional magnetic resonance imaging, ECoG: electrocorticography, LFP: local field potential, MEA: micro-electrode array, ME: microelectrode, <i>blue</i> color: non-invasive methods, <i>red</i> color: invasive methods) [7].	6
Figure 2-3 : A portable EEG system with a cap on which the electrodes are placed, a biopotential amplifier, and a recording/monitoring device [10].	7
Figure 2-4 : (a) Patient undergoing an MEG. (b) Entrance to magnetically shielded MEG room [17].	8
Figure 2-5 : (a) ECoG electrodes over the cortex [14]. (b) Cortical microelectrode array [15].	9
Figure 2-6 : Grand average of raw P300 target signal (80 epochs) and non-target signal (1120 epochs). The thick lines represent the mean, and thin lines represent the mean plus and minus the standard deviation [18].	11
Figure 2-7 : The EEG signal spectrum in response to a visual stimulation with a flickering frequency of 7 Hz. The response is observed as the peaks at 7 Hz and its harmonics [19].	12
Figure 2-8 : A 6 by 6 P300 speller matrix. Third row is intensified [27].	16
Figure 2-9 : P300 paradigms used for locomotion [18], [34].	18
Figure 3-1 : Separating hyperplanes possible to be selected as classifier for the problem	30

Figure 3-2 : Optimum separating hyperplane maximizes the margin width determined by the support vectors.....	31
Figure 3-3 : The objective of the SVM algorithm is to maximize the margin width and minimize the distance of the misclassified observations to their class boundary.	32
Figure 3-4 : Observations separable in 1D	32
Figure 3-5 : Observations not separable in 1D	33
Figure 3-6 : Observations that are not separable in 1D being separable in 2D. ...	33
Figure 3-7 : Three layer feed-forward artificial neural network.....	34
Figure 3-8 : ANN piece	36
Figure 4-1 : The placement of the electrodes in the experiment	42
Figure 4-2 : RMS Error calculated at each iteration for subject 1 under linear feature normalization and unit norm feature vector normalization with PCA-Coefficient 99.5.	51
Figure 4-3 : Validation accuracy calculated at each iteration for subject 1 under linear feature normalization and unit norm feature vector normalization with PCA-Coefficient 99.5.	51
Figure 4-4 : The timing scheme for a single trial of the experiment [69].....	60
Figure 4-5 : Electrode configuration used in the experiment [69].....	60
Figure 4-6 : The time segments used in feature extraction for training and evaluation data (adapted from [69]).....	61
Figure 4-7 : The images representing (a) tongue, (b) left hand, (c) right hand, (d) left foot, (e) right foot movement imagination and (f) the fixation cross used in the trials.	67
Figure 4-8 : Electrode configuration used in the experiment.	68
Figure 4-9 : The first level of the menu. The elements are selected by performing the tasks symbolized with small icons at the top left corner of each element.	78
Figure 4-10 The second level of the menu. The commands for (a) motorized bed, (b) TV, (c) air conditioner, (d) light bulb are selected by performing the tasks symbolized with small icons at the top left corner of each element.	78

Figure 4-11 : The interface of the test application implemented. The EEG dataflow is checked before and after each run from the window in upper right corner..... 80

Figure 4-12 : The time scheme of a single trial for N repetitions..... 81

CHAPTER 1

INTRODUCTION

Interaction with the outside world is one of the main existence reasons of the human being. However this interaction cannot be realized by all the people. The normal pathways to sense and express may be lost or damaged due to some accidents or diseases. For these people, also called as locked-in, Brain Computer Interfaces (BCIs) play an important role in terms of providing alternative pathways to interact with the outside world. For that purpose, researchers from many fields have been working in this area to improve the quality of life of these people. The researches mainly focus on enhancing the accuracy and speed of these systems by improving the signal acquisition, understanding neurophysiological activity of the brain and optimizing the signal processing techniques used in the system.

Several non-invasive and invasive signal acquisition techniques have been used in BCI research. In non-invasive electroencephalography (EEG) and magnetoencephalography (MEG), the electromagnetic activity of the brain is measured by the electrodes placed over the skull. In invasive electrocorticography (ECoG), single micro-electrode (ME), micro-electrode array (MEA), and local field potentials (LFPs), the electrodes are placed surgically inside the skull to measure the cortical activity. Functional Magnetic Resonance Imaging (fMRI) and Near Infrared Spectroscopy (NIRS), in which regional changes in cerebral blood oxygenation levels are detected non-invasively, are also used in BCI

research [1], [2]. Among these techniques EEG is preferred in this study, due to the hardware's low cost, low risk and portability. Also its temporal resolution is sufficiently high for online BCI applications [3].

There are various neurophysiological mechanisms that have been shown to be useful for BCI applications. These mechanisms may be either the response of the brain to an event or the activity generated by the subject independent from an event. Among these mechanisms sensorimotor rhythms (SMR), which are generated during motor imagery, are studied in this study. In BCI applications, SMR are converted to control signals that enable interaction with the outside world. Since SMR do not require any visual or auditory stimuli, they are widely used in BCI applications [4].

In order to convert SMR to BCI control signals, several signal processing techniques have been used in the literature. These techniques can be analyzed in three steps. First one is the signal enhancement step. In that part the quality of signal is improved by applying techniques like filtering, down-sampling, etc. Second step is feature extraction. In that part, the relevant information for the application is obtained from the data. The final step is the classification in which a mathematical model is constructed using the normalized features extracted in the previous step. The constructed model is used to produce control signals related to the application. Specifically in this study, band-pass filtering, Laplacian filtering, and common average reference (CAR) filtering are used to enhance the EEG signal. For feature extraction; Common Spatial Pattern (CSP), Power Spectral Density (PSD), and Principal Component Analysis (PCA) are tested. Linear Feature Normalization (LFN), Gaussian Feature Normalization (GFN), and Unit-norm Feature Vector Normalization (UFVN) are studied in Support Vector Machine (SVM) and Artificial Neural Network (ANN) classification. The performances of these methods are first evaluated in the datasets provided for BCI Competition III and IV. Then, different tasks have been studied in the METU

Brain Research Laboratory to find the optimum number and type of tasks to control a motor imagery based BCI.

In this study, a SMR based BCI that assist the paralyzed people for controlling environmental devices is designed and a test application for the design is realized. The aim of the design is basically make it possible for a subject to select items from a menu by the help of motor imagery tasks.

The thesis starts with an introductory chapter presenting the BCI research in the literature with the analysis of its building blocks (CHAPTER 2). Then, in CHAPTER 3, the signal processing techniques in BCI research will be analyzed in detail focusing on the methods used in the study. CHAPTER 4 provides the results of the experiments conducted in this study. Finally, in CHAPTER 5, the study is summarized and the conclusions on the results is provided.

CHAPTER 2

BRAIN COMPUTER INTERFACES

Brain Computer Interfaces (BCIs) are the systems that convert brain signals into control signals that are necessary to interact with the outside world. These systems may be the only or preferred pathway for

- “the patients with severe motor disabilities who lost voluntary muscle control,
- the patients with Amyotrophic Lateral Sclerosis (ALS) who has to accept artificial ventilation to prolong life as the disease progresses,
- children and adults with severe cerebral palsy who do not have useful muscle control,
- patients with brainstem strokes who have only minimal eye movement control,
- individuals with severe muscular dystrophies or peripheral neuropathies,
- people with acute disorders causing extensive paralysis (e.g., Landry-Guillain- Barré syndrome),
- patients with high cervical spinal cord injuries” [5].

Therefore, researchers from many fields have been working in this topic to improve the quality of life of these people.

Essential elements and operation of a typical BCI are given in Figure 2-1. These elements will be analyzed in detail, step by step in this chapter. At first, information about the signal acquisition techniques used in BCI applications will

be given. Then the activity of the brain will be analyzed considering the role of that activity in a possible BCI application. In the following section the signal processing stage in a typical BCI will be explained briefly. Finally the BCI applications in the literature will be reviewed.

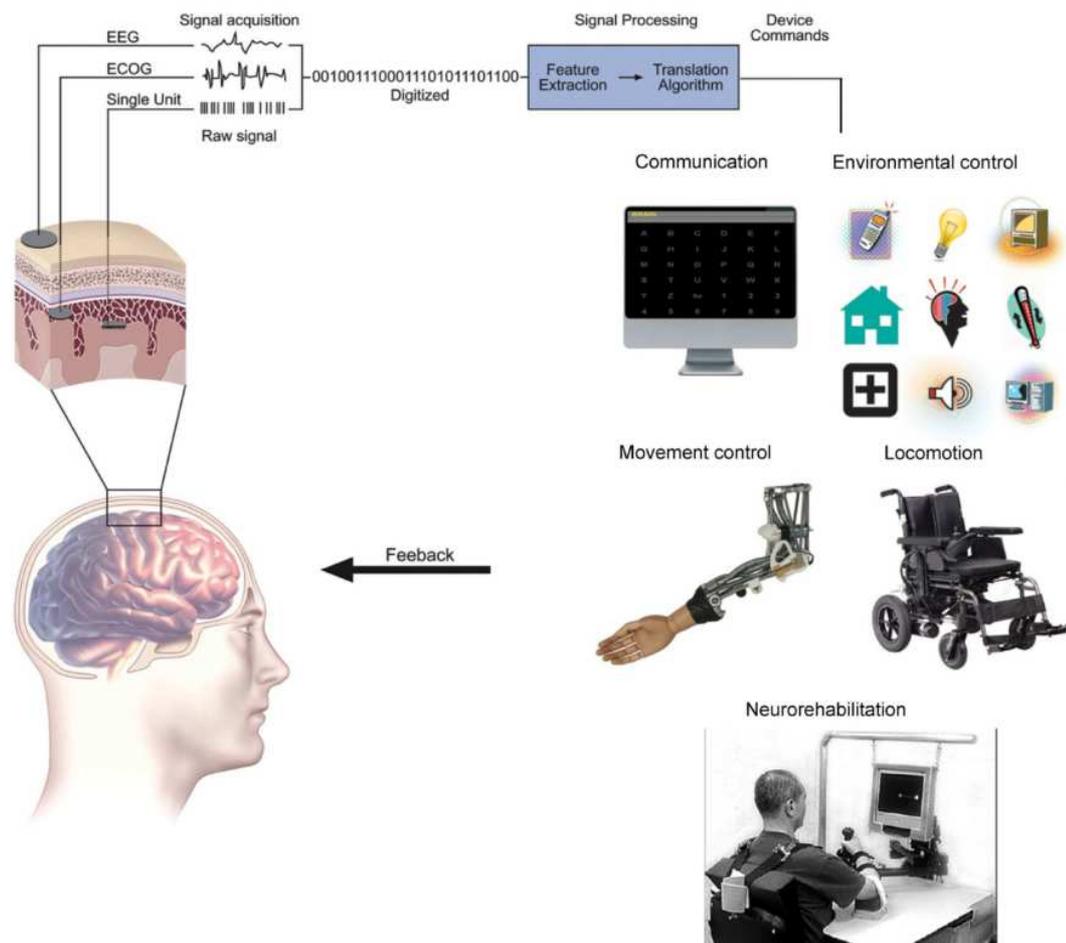


Figure 2-1 : The brain computer interface cycle [6].

2.1 Signal Acquisition

There are several non-invasive and invasive signal acquisition techniques used in BCI research. In non-invasive electroencephalography (EEG) and magnetoencephalography (MEG), the electromagnetic activity of the brain is

measured by the electrodes placed over the skull. In invasive electrocorticography (ECoG), single micro-electrode (ME), micro-electrode array (MEA), and local field potentials (LFPs), the electrodes are placed surgically inside the skull to measure the cortical activity. Functional Magnetic Resonance Imaging (fMRI) and Near Infrared Spectroscopy (NIRS), in which regional changes in cerebral blood oxygenation levels are detected non-invasively, are also used in BCI research [1], [2]. A schematic overview of the signal acquisition techniques used in BCI research is given in Figure 2-2.

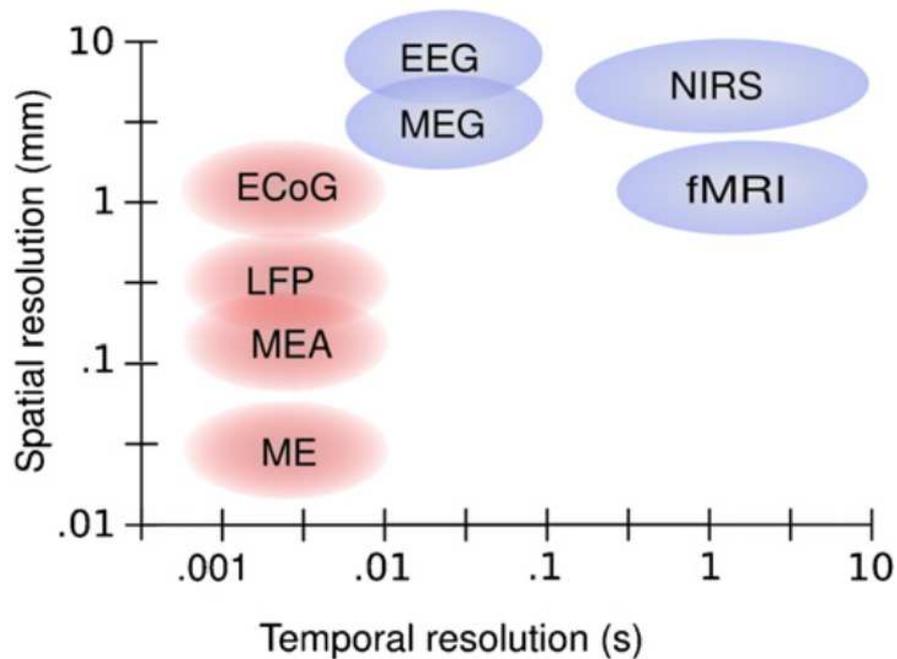


Figure 2-2 : The spatial and temporal resolution scale of the signal acquisition techniques used in the BCI study (EEG: electroencephalography, MEG: magnetoencephalography, NIRS: near-infrared spectroscopy, fMRI: functional magnetic resonance imaging, ECoG: electrocorticography, LFP: local field potential, MEA: micro-electrode array, ME: microelectrode, blue color: non-invasive methods, red color: invasive methods) [7].

A closer look inside EEG, ECoG, ME, MEA, LFPs and MEG will be taken in the following subsections, since they have sufficient temporal resolution for a real time BCI.

2.1.1 Electroencephalography (EEG)

Electroencephalography (EEG) is the recording of electrical activity within the brain using the electrodes placed over the skull [8], [9]. It is invented by Hans Berger in 1929. The noisy and low amplitude signal (of the order of 10^{-4} Volts) is filtered and amplified considering the frequency characteristic of the signal to be detected and noise (i.e. 50 Hz supply noise) to be suppressed. Then, the signal is recorded after being digitalized (Figure 2-3).



Figure 2-3 : A portable EEG system with a cap on which the electrodes are placed, a biopotential amplifier, and a recording/monitoring device [10].

EEG has a temporal resolution sufficiently high for online applications. On the other hand, its spatial resolution is low due to the blurring effect of the head tissue. Also, the measured EEG signal may contain artifacts originating from the movement of the electrodes, eye blink or muscular activity. Furthermore, in most of the EEG devices, the electrodes are placed on the skull by applying conductive gel in order to decrease the contact impedance. This is also a disadvantage in

terms of practical use. Even though, EEG is preferred in most of the BCI studies due to its low cost, low risk and portability [3]. For a selection of EEG based BCI studies, the reader may refer to [2, 10-12].

2.1.2 Magnetoencephalography (MEG)

Magnetoencephalography is a non-invasive technique for measuring the tiny magnetic field fluctuations (about 10^{-14} Tesla) induced by the populations of cerebral neurons. Its temporal resolution is comparable to that of EEG [3].

Since the MEG signals have low amplitude, the measurements must be performed in a magnetically shielded room (Figure 2-4b) to avoid the signals being distorted. Also, it usually requires a large cooling unit for its sensors. Therefore, MEG systems are rather expensive and non-portable (Figure 2-4a) [3].

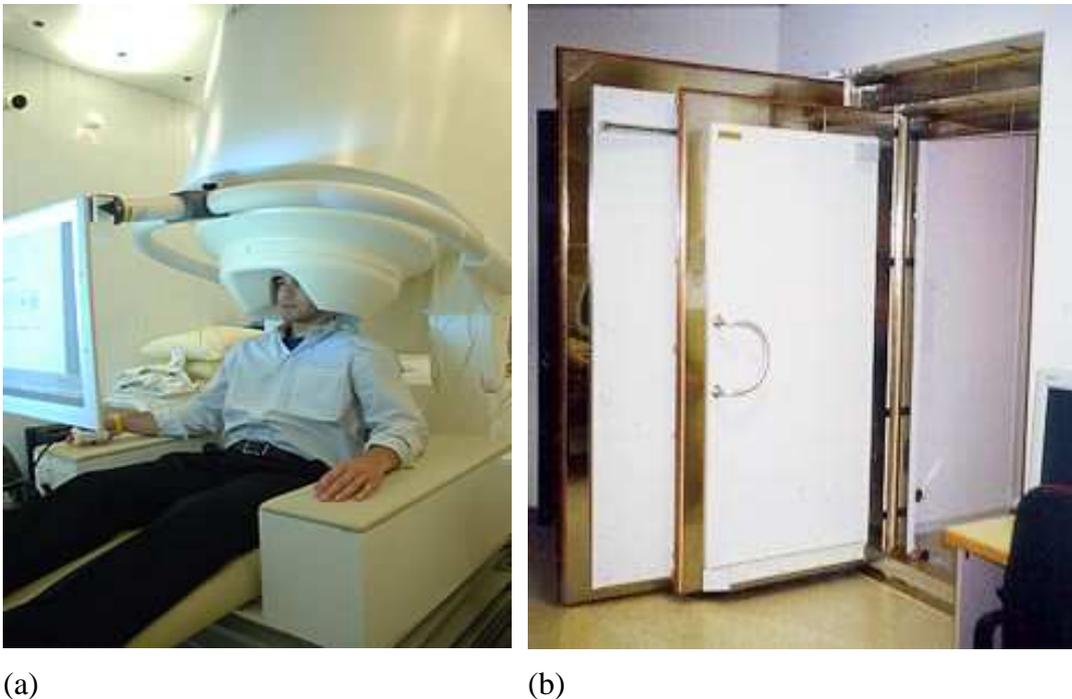


Figure 2-4 : (a) Patient undergoing an MEG. (b) Entrance to magnetically shielded MEG room [17].

2.1.3 Other Signal Acquisition Techniques

There are also other signal acquisition techniques used in the BCI research. Among these, Electrocorticography (ECoG) (Figure 2-5a), micro-electrode (ME), micro-electrode array (MEA) (Figure 2-5b), and local field potentials (LFPs) are invasive signal acquisition techniques in which the brain signals are measured by the help of electrodes placed surgically inside the skull. [7].

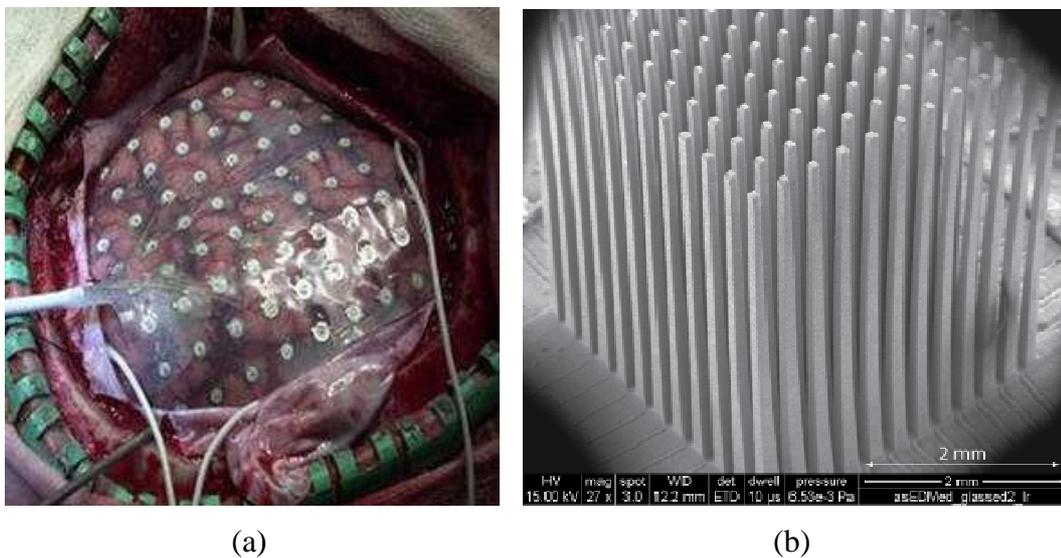


Figure 2-5 : (a) ECoG electrodes over the cortex [14]. (b) Cortical microelectrode array [15]

These techniques have higher signal-to-noise ratio and spatial resolution compared to EEG [3]. Therefore they play an important role in brain research including BCI [11, 12]. However, they are not widespread due to the risk in the surgical operation performed.

2.2 Neurophysiologic Signals

There are various neurophysiological mechanisms that have shown to be useful for BCI applications. These mechanisms are categorized into *evoked* and *induced responses* considering their dependency on a stimulus. Evoked potentials are the response of the brain to an event. Therefore the synchronization of the event and the EEG signal is important for evoked responses. Induced responses are generated by the subject independent from an event.

2.2.1 Evoked Responses

Evoked responses widely used in BCI applications are P300, Steady-State Visual Evoked Potential (SSVEP), and Slow Cortical Potentials (SCPs). In the following subsections, these responses will be analyzed in detail.

2.2.1.1 P300

P300 is a peak that typically occurs 300 ms after an expected, but infrequent, random event occurs (Figure 2-2). It is a natural neuromechanism that almost all subjects have without requiring any training period. In P300 based BCI applications, each stimulus event corresponds to a symbol/picture with a particular meaning for the interface (e.g. letters, high level commands). Among these symbols, the target of the subject is detected depending on the P300 peak occurrence time.

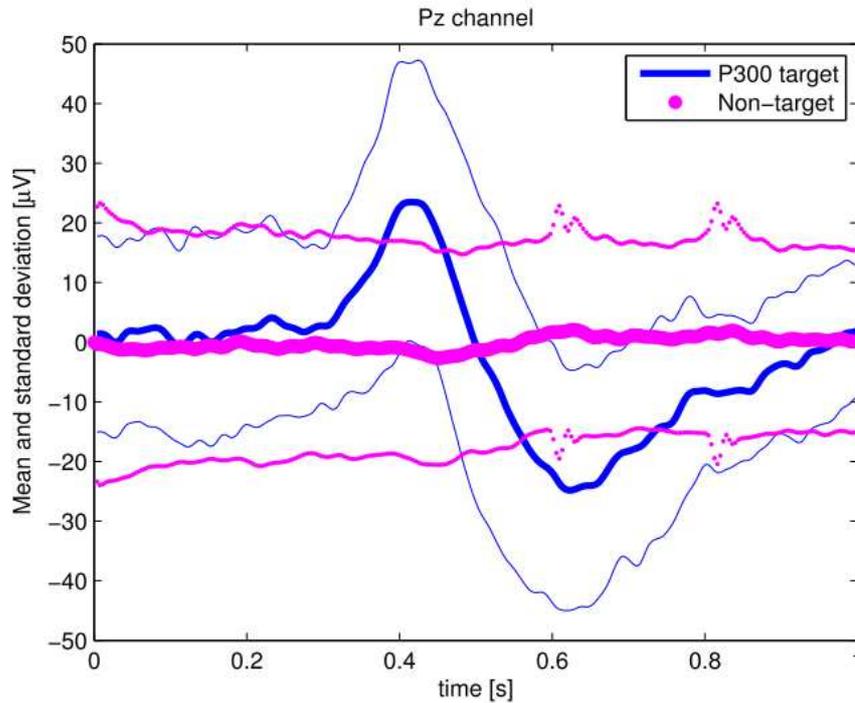


Figure 2-6 : Grand average of raw P300 target signal (80 epochs) and non- target signal (1120 epochs). The thick lines represent the mean, and thin lines represent the mean plus and minus the standard deviation [18].

If a visual stimulus is used in the application, it must be perceptible on the user field of view without gazing the specific stimulus. Another disadvantage of P300 arrives from the fact that the user has to wait for the occurrence of the desired (target) stimulus which randomly appears. It is not the user who decides when to provide an intention but rather the emergence of the stimulus. Moreover, processing algorithms have to run synchronously with the start of the stimuli. Also, increasing the number of possible commands (events) decreases the transfer rate because each stimulus is flashed less frequently [18].

2.2.1.2 Steady-State Visual Evoked Potential (SSVEP)

When a stimulus flickering at a constant frequency greater than 5Hz is presented to a subject, a potential response at the same frequency and its harmonics is

observed at the occipital brain region (visual cortex) of the subject (Figure 2-7). This response is called Steady-State Visual Evoked Potential (SSVEP) [19].

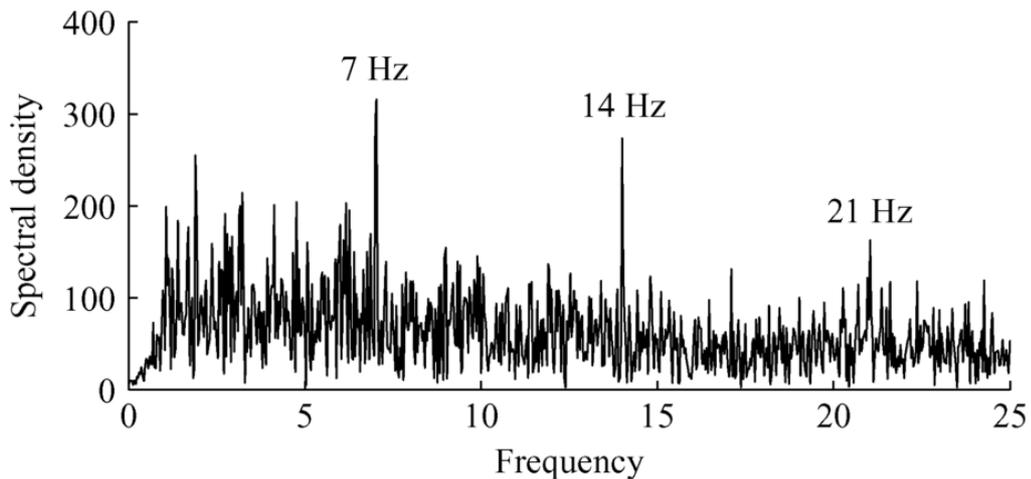


Figure 2-7 : The EEG signal spectrum in response to a visual stimulation with a flickering frequency of 7 Hz. The response is observed as the peaks at 7 Hz and its harmonics [19].

In SSVEP based BCI applications, the user has to gaze the stimuli (representing an action, letter, etc.) positioned in some part of the screen which involves the movement of the eyes. Because it depends on the brain's normal output pathway of peripheral nerves and eye muscles it cannot be called a true BCI. Notwithstanding this, the interface can be suitable for people with severe motor disabilities but still able to perform small eye movements [18].

2.2.1.3 Slow Cortical Potentials (SCPs)

Slow cortical potentials are the negative or positive polarizations of the electromagnetic activity in the brain. These potentials are generated by the subject voluntarily. SCP based BCI applications may require extensive training periods depending on the subject's ability to shift her/his SCP. Also, the modulation of

SCPs is relatively slow. Therefore, the amount of information transmitted per unit time is limited in SCP-based BCI applications [20].

2.2.2 Induced Responses

Induced responses widely used in BCI applications are sensorimotor activity and responses to mental tasks. In the following subsections, these responses will be analyzed in detail.

2.2.2.1 Sensorimotor Rhythms (SMR)

One way of describing the brain signals is to divide the rhythmic activity into frequency bands. These bands are the delta (δ) band, the theta (θ) band, the alpha (α) band, the beta (β) band, the gamma (γ) band and the mu (μ) band. The frequency range, related brain region and the mental state they appear are given in [9] in detail. Among these, μ and β bands (8-12 Hz & 13-30 Hz respectively) are related to sensorimotor activity and they are widely used in BCI applications.

μ and β rhythms originate in the primary sensorimotor cortex. A voluntary movement results in a desynchronization in the μ and β bands (*event related desynchronization, ERD*). After the movement, the power in the brain rhythm increases (*event related resynchronization, ERS*) [21]. Similar to real movement execution case, imagination of movement can also modify the neuronal activity in the sensorimotor cortex. This phenomenon makes it possible for patients with severe motor disabilities to use BCI with motor imagery. In these BCI applications the type of motor imagery (right/left hand/foot movement) is identified by classifying the power in the μ and β bands at electrodes located over the primary sensorimotor cortex [22].

The major disadvantage in ERD/ERS based BCIs is the long training period. It may take a few months depending on the subject's ability to control her/his sensorimotor rhythms.

2.2.2.2 Responses to Mental Tasks

In addition to the motor imagery, different non-movement mental tasks (e.g., solving a multiplication problem, mental counting, imagining a 3D object) can be used in BCI systems. Each task has a specific distribution of EEG frequency pattern over the scalp [21]. Compared to motor imagery, these tasks are more complicated. Therefore, they are not widely used in practical BCI applications.

2.3 Signal Processing

Signal processing in BCI applications can be analyzed in three steps. First one is the *signal enhancement* step. In that part the quality of signal is improved by applying techniques like filtering, down-sampling, etc. Second step is *feature extraction*. In that part the relevant information for the application is obtained from the data. The final step is the *classification* in which a mathematical model is constructed using the normalized features extracted in the previous step. The constructed model is used to produce control signals related to the application. A review of the signal processing methods used in BCI applications and the detailed information about the methods used in the study is given in CHAPTER 3.

2.4 BCI Applications

In this section BCI applications in the literature will be reviewed by dividing into groups as applications for communication, environmental control, movement control, locomotion, and neurorehabilitation.

2.4.1 Communication

BCIs for communication focus on selection of *icons*. These icons vary from low level (i.e. letters) to high level (i.e. words, sentences). Selection of these icons has been realized depending on different neurophysiologic mechanisms. In [23] binary selection among letter-banks is performed using SCPs. The letter groups are split into two until a single letter remains. Movement of cursor in 1D or 2D is also utilized to select icons. In [24], SMR is used to move cursor in 2D. In the experiments with different patients, the best performance is achieved with 92% hit rate, 1.9 seconds movement time, and 4.9 movement precision (target size as % of workspace). Among all approaches, P300 based BCI for communication has been the most popular one. In the spelling paradigm proposed by Farwell and Donchin [25], a 6 by 6 matrix of characters is presented to the subject on a computer screen (see Figure 2-8). The rows and columns of this matrix are intensified sequentially in a random order. When the row or column containing the target character is intensified, P300 potential is evoked. Therefore, after a few repetitions the target character is determined using the instant at which P300 potential is evoked. In [26], which is one of the latest works on P300 spelling paradigm, one character per 9.6 sec. speed is obtained with 94.5% accuracy.



Figure 2-8 : A 6 by 6 P300 speller matrix. Third row is intensified [27].

The reader can find detailed information about current silent speech methodologies for normal and disabled individuals in [13].

2.4.2 Environmental Control

Controlling devices like air conditioner, power bed, TV, light etc. could greatly improve the quality of life of a patient with severe motor disabilities. A pilot study in which a system was implemented and validated to allow disabled persons to improve their mobility and communication within the surrounding environment is reported in [28]. In the integrated framework developed, keyboard, mouse, joystick, trackball touchpad, buttons, microphone, and head tracker were also utilized. When the user was not able to use any of these devices, a BCI is suggested by the support team. Using a SMR based BCI, average accuracy higher than 75% (accuracy expected by chance alone was 50%) is obtained in a binary selection task among the icons.

2.4.3 Movement Control

There are many researches on restoration of motor control with robotic and prosthetic devices in paralyzed patients. These researches can be divided into two groups as invasive and non-invasive.

Invasive methods are mostly applied on animals like rats [29] and monkeys [30]. In one of these animal experiments, [30], intracortical microelectrodes were implanted in the proximal arm region of the primary motor cortex of a monkey. In the experiments the monkey is trained to feed itself with a robotic arm moving in 3D and 61% success rate (the percentage of attempted trials where the monkey succeeded in getting the food into its mouth) is obtained. In a recent pilot study, a 96-microelectrode array is implanted in primary motor cortex of a tetraplegic human to measure neuronal activity. In the study, the patient achieved to open and close a prosthetic hand, and to perform rudimentary actions with a multi-jointed robotic arm.

Non-invasive researches mainly focus on EEG recordings over the sensorimotor cortex. In a SMR based application, [31], a tetraplegic patient achieved to control the opening and closing of his normally paralyzed hand by an orthosis using the motor imagery of two limbs (e.g left vs. right hand or right hand vs. both feet) with nearly 100% accuracy after 5 months of training period. In another motor imagery based study, Gernot Müller-Putz and his colleagues performed experiments for 3 days with a patient with a spinal cord lesion to control an implanted neuroprosthesis. In the study, the patient used his EEG to step through several phases of a hand grip with 73% best performance [32]. In a recent SSVEP based BCI study, seven subjects performed two tasks: moving orthosis through four different positions and grasping object. Although none of the subjects had any training, six subjects showed good control with a performance higher than 60% [33].

2.4.4 Locomotion

Many researches have been done on developing a BCI-driven wheelchair in order to provide mobility to the patients. In the EEG based BCI-driven wheelchair simulations, Galán and his colleagues used left hand movement imagination to turn left, rest to go forward, and word association to turn right. They also used the information provided by the wheelchair's sensors. In the experiments two subjects were able to reach 100% (subject 1) and 80% (subject 2) of the final goals along the pre-specified trajectory in their best sessions by delivering a mental command at every 0.5 sec. Pires and Nunes developed a P300 paradigm (Figure 2-9a) to control wheelchair through specific directions and performed some offline tests [18]. Palankar and his colleagues used a similar P300 paradigm (Figure 2-9b) for real-time control of a wheelchair-mounted robotic arm. In the system developed, 1 output at every 15 sec. time duration is produced with a performance higher than 80% to direct the robot along a step-by-step path to a desired position [34].

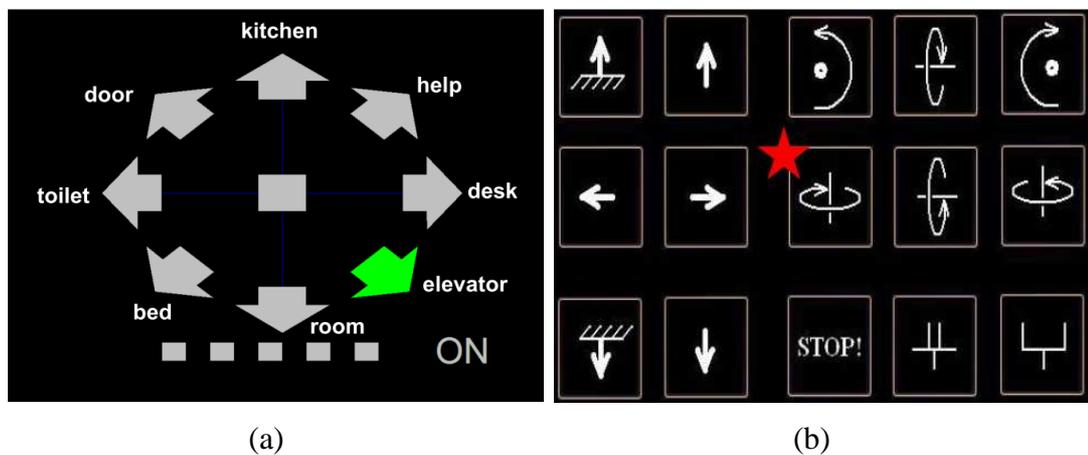


Figure 2-9 : P300 paradigms used for locomotion [18], [34].

2.4.5 Neurorehabilitation

The aim of BCI applications mentioned up to that point was to improve the quality of life of the patient with assistive devices. The aim of the BCI applications for neurorehabilitation is to help the patient restore motor function after stroke or in other chronic central nervous system (CNS) traumatic injuries or disease. There are two strategies for that purpose. In the first strategy the patient is trained to produce more normal brain activity which is measured by specific EEG features. In the second strategy, the patient uses a device that assists movement depending on the brain activity. This strategy improves motor function yielding sensory input that induces CNS plasticity. The reader can find detailed information about the BCIs in neurological rehabilitation in [5].

CHAPTER 3

SIGNAL PROCESSING IN BRAIN COMPUTER INTERFACES

In section 2.3, a brief description about signal processing procedure in a typical BCI system is given. In this chapter, the signal processing techniques in BCI research will be analyzed in detail focusing on the methods used in the study. These methods can be summarized as follows:

- For signal enhancement: band-pass filtering, Laplacian filtering, and common average reference (CAR) filtering
- For feature extraction: Common Spatial Pattern (CSP), Power Spectral Density (PSD), Principal Component Analysis (PCA)
- For normalization: Linear Feature Normalization (LFN), Gaussian Feature Normalization (GFN), Unit-norm Feature Vector Normalization (UFVN)
- For classification: Support Vector Machines (SVM), Artificial Neural Networks (ANN)
- For evaluation: classification accuracy, Cohen's kappa coefficient, Nykopp's information transfer

3.1 Signal Enhancement

As it is mentioned in 2.1.1, EEG has relatively low signal-to-noise ratio. The reason is that the signals are spatially blurred due to volume conduction in the intervening tissue. Also, the measured EEG signal may contain artifacts originating from the movement of the electrodes, eye blink or muscular activity. Therefore, signal enhancement plays an important role in the analysis of the EEG data.

The most common types of signal enhancement techniques are artifact detection, spectral filtering and spatial filtering. Artifact detection attempts to find confounding signals from sources outside the brain, such as eye and muscle artifacts, and then attempts to remove them. Spectral filtering is used to remove noise signals (e.g. line noise) and select the frequency band related to the activity of the brain in the application. Spatial filtering also improves the signal-to-noise ratio of the signal by re-referencing the EEG channels [7].

In this study, the signal enhancement is provided by spatial and spectral filters. As spectral filter, a Butterworth filter is designed using MATLAB Filter Design Toolbox. The EEG data is band-pass filtered between 8-30Hz, which is the band (μ and β rhythms; 8-12 Hz and 13-30 Hz, respectively) related to sensorimotor activity [21]. As spatial filter, four different re-referencing techniques are utilized. These are standard ear-reference, a common average reference (CAR), a small Laplacian (SL) and a large Laplacian (LL). In the standard ear-reference technique, all the electrodes are directly referred to the electrodes placed on the ears. In the CAR, the entire average of the potentials at the channels is subtracted from of each channel. In the Laplacian filtering, the weighted sum of the voltage in the surrounding electrodes is subtracted from that of the voltage in the channel of interest. The calculation of the Laplacian is performed according formula given below.

$$V_i^{LAP} = V_i^{ER} - \sum_{j \in S_i} g_{ij} V_j^{ER} \quad (3-1)$$

where

$$g_{ij} = \frac{1/d_{ij}}{\sum_{j \in S_i} 1/d_{ij}} \quad (3-2)$$

V_i^{ER} is the ear-referenced voltage. S_i is the set of electrodes surrounding the i^{th} electrode, and d_{ij} is the distance between electrodes i and j (where j is a member of S_i). For the small Laplacian, S_i is the set of neighbor electrodes approximately 3 cm to the center electrode. For the large Laplacian, it is the set of neighbor electrodes approximately 6 cm to the center electrode [35].

3.2 Feature Extraction

Feature extraction is summarizing the measurements for a classification problem while still describing the data with sufficient accuracy. In feature extraction process, high dimensional and possibly redundant data is transformed into a reduced representation set of feature vectors.

In BCI applications several feature extraction methods are utilized depending on the pattern worked on. For a P300 based application, simply a voltage threshold at a specific time instant may be a feature. However, in an SMR-based application the frequency characteristic of the signal in a specific band may be more important. Also in other applications, autoregressive (AR) and adaptive autoregressive (AAR) parameters [36,37], time-frequency features [38], and inverse model-based features are used [39-41].

In this study Common Spatial Pattern (CSP), Power Spectral Density (PSD), and Principal Component Analysis are used to extract feature from the motor imagery

related EEG signal. CSP, which analyzes the signal in time domain, is one of the most popular feature extraction algorithms in the BCI study [42-44]. PSD analyzes the signal in frequency domain and it is also a common method in SMR-based BCI applications [45-48]. PCA is a classical dimension reduction technique, which has applications in the BCI research [45,49,50]. In the following subsections these three methods will be analyzed in detail.

3.2.1 Common Spatial Pattern (CSP)

The goal of common spatial pattern (CSP) analysis is to design spatial filters so that the filtered time series have variances optimal for the discrimination. In this section, the method will be explained for a 2-class problem. However, the method can be extended to multiclass applications [51].

Let the raw EEG data of a single trial is represented by an $N \times T$ matrix \mathbf{E} , where N is the number of channels and T is the number of samples. The normalized spatial covariance of the EEG is calculated as follows.

$$\mathbf{C} = \frac{\mathbf{E}\mathbf{E}'}{\text{trace}(\mathbf{E}\mathbf{E}')} \quad (3-3)$$

where $'$ is the transpose operator and $\text{trace}(\mathbf{x})$ is the sum of the diagonal elements of \mathbf{x} . For each of the two distributions to be separated, the spatial covariances, $\overline{\mathbf{C}}_1$ and $\overline{\mathbf{C}}_2$, are calculated by averaging over the trials of each group. The composite spatial covariance is calculated as

$$\mathbf{C}_c = \overline{\mathbf{C}}_1 + \overline{\mathbf{C}}_2 \quad (3-4)$$

Then \mathbf{C}_c is factored as $\mathbf{C}_c = \mathbf{U}_c \boldsymbol{\lambda}_c \mathbf{U}_c'$ where \mathbf{U}_c represents the matrix of eigenvectors and $\boldsymbol{\lambda}_c$ represents the matrix of eigenvalues which are sorted in

descending order at the diagonal. Afterwards, the whitening transformation \mathbf{P} is calculated as follows;

$$\mathbf{P} = \sqrt{\lambda_c^{-1}} \mathbf{U}_c' \quad (3-5)$$

This transformation (\mathbf{P}) equalizes the variances in the space spanned by \mathbf{U}_c , i.e., all eigenvalues of $\mathbf{P}\mathbf{C}_c\mathbf{P}'$ are equal to one. If $\overline{\mathbf{C}}_1$ and $\overline{\mathbf{C}}_2$ are transformed as

$$\mathbf{S}_1 = \mathbf{P}\overline{\mathbf{C}}_1\mathbf{P}' \text{ and } \mathbf{S}_2 = \mathbf{P}\overline{\mathbf{C}}_2\mathbf{P}' \quad (3-6)$$

then \mathbf{S}_1 and \mathbf{S}_2 share the same eigenvectors, i.e.,

$$\text{if } \mathbf{S}_1 = \mathbf{B}\lambda_1\mathbf{B}' \text{ then } \mathbf{S}_2 = \mathbf{B}\lambda_2\mathbf{B}' \text{ and } \lambda_1 + \lambda_2 = \mathbf{I} \quad (3-7)$$

Here, \mathbf{I} represents the identity matrix which means the sum of two corresponding eigenvalues is always one. Therefore, the eigenvector corresponding to the largest eigenvalue of \mathbf{S}_1 has the smallest eigenvalue for \mathbf{S}_2 and vice versa. Therefore, the eigenvectors \mathbf{B} becomes useful in classifying two distributions. When the EEG data is whitened with \mathbf{P} and projected onto a group of eigenvectors in \mathbf{B} , it becomes optimal for separation of two populations in the least squares sense.

With the projection matrix $\mathbf{W} = (\mathbf{B}'\mathbf{P})'$, the EEG data is filtered as

$$\mathbf{Z} = \mathbf{W}\mathbf{E} \quad (3-8)$$

The columns of \mathbf{W}^{-1} are referred as common spatial patterns. They represent time-invariant EEG source distribution vectors. The m first and last rows of \mathbf{Z} ($\mathbf{Z}_p : p = 1, \dots, 2m$) are used while calculating the final feature vectors as follows;

$$f_p = \log \left(\frac{\text{var}(\mathbf{Z}_p)}{\sum_{i=1}^{2m} \text{var}(\mathbf{Z}_i)} \right) \quad (3-9)$$

The feature vectors \mathbf{f}_p are used to train a classifier. The log-transformation is due to approximate normal distribution of the data [52].

In this study, the multiclass extension of CSP is realized by calculating \mathbf{W} for each two of the N classes, and concatenating the resultant \mathbf{f} vectors. For an N -class problem, there exists $\binom{N}{2}$ different projection matrices, \mathbf{W} . Each projection yields a p -dimensional feature vector \mathbf{f} . These vectors are concatenated and final feature vector of size $p \binom{N}{2}$ is obtained. This multiclass extension methodology is also utilized by one of the winner algorithms in BCI Competition IV [53].

3.2.2 Power Spectral Density (PSD)

Power Spectral Density (PSD) is a positive real function which describes the power distribution of a signal over frequency. There are several parametric and non-parametric approaches for estimation of this distribution. In a common parametric technique, an autoregressive model is fitted to the observations. A common non-parametric technique is the Welch's periodogram method which is also used in the calculation of PSD features classified in this study [54].

In Welch's periodogram method, the signal is split up into overlapping segments. Then, squared magnitude of discrete Fourier transform of each segment is calculated after windowing. The final PSD is estimated taking the average of PSD estimation of each segment [55].

3.2.3 Principal Component Analysis (PCA)

Principal component analysis (PCA) is a classical statistical method which is invented by Pearson K. in 1901 [56]. This linear transform has been widely used in data analysis and compression to convert a set of observations of possibly correlated variables into a set of values of uncorrelated variables called principal components. In this study PCA is used to extract feature from the PSD data. By using PCA, PSD data is projected from the higher d-dimensional space to the lower k-dimensional eigenspace, which is composed of k eigenvectors, and retain the feature information. This reduction in dimension may improve the performance and classification accuracy of the classifier used in the problem. The calculation of PCA is given step by step below.

Step 1: Compute the mean vector

$$\mathbf{m} = \frac{1}{nTr} \sum_{i=1}^{nTr} \mathbf{p}_i \quad (3-10)$$

where $\mathbf{p}_i = [p_1 \dots p_d]^T$ is the i^{th} d-dimensional training sample (i.e., PSD value of all channels) and nTr is the number of the training samples.

Step 2: Calculate the covariance matrix

$$\mathbf{\Sigma} = \sum_{i=1}^{nTr} (\mathbf{p}_i - \mathbf{m})(\mathbf{p}_i - \mathbf{m})^T \quad (3-11)$$

where $\mathbf{\Sigma}$ is a dxd matrix.

Step 3: Find the eigenvectors and corresponding eigenvalues of $\mathbf{\Sigma}$ by solving

$$\mathbf{\Sigma} \mathbf{x} = \lambda \mathbf{x} \quad (3-12)$$

Let the solution vectors $\mathbf{x} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_d\}$ represent the eigenvectors, and $\boldsymbol{\lambda} = \{\lambda_1, \lambda_2, \dots, \lambda_d\}$ represent the corresponding eigenvalues in the descending order. \mathbf{e}_1 is called the *principle component* of the dataset and it represents the most significant data dimension.

Step 4: Generate a $d \times k$ matrix A whose columns consist of the k eigenvectors corresponding to the largest eigenvalues:

$$A = [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k] \quad (3-13)$$

Step 5: Represent the data in k -dimensional subspace by performing the projection operation;

$$\mathbf{p}' = A^t(\mathbf{p} - \mathbf{m}) \quad (3-14)$$

where $\mathbf{p}' = [\mathbf{p}'_1 \dots \mathbf{p}'_k]^T$ is the PCA feature vector.

When the PCA feature of the testing data is extracted, only the operation in Step 5 is performed by replacing the training data \mathbf{p} with the testing data [45].

The ratio of the sum of the eigenvalues $\{\lambda_1, \lambda_2, \dots, \lambda_k\}$ to the sum of all eigenvalues $\{\lambda_1, \lambda_2, \dots, \lambda_d\}$, represents the information in the space spanned by the corresponding eigenvectors. This ratio will be referred as PCA-coefficient and will be used in PCA-based feature extraction experiments performed in this study.

3.3 Normalization

Normalization is an important step in terms of the accuracy in the classification stage. If the ranges of the features are unbalanced, they may have different importance while classifying the data. In order to make the role of the features in classification independent from the range of the features, they must be normalized identically. One idea for feature normalization is to set the mean of each feature to zero, and the variance to one [57]. This is called Gaussian normalization which can be formulized as follows:

$$\hat{\mathbf{x}} = \frac{\mathbf{x} - \overline{\overline{\mu_x}}}{\overline{\overline{\sigma_x^2}}} \quad (3-15)$$

where $\overline{\mu_x}$ and $\overline{\sigma_x^2}$ are the mean and variance vectors of the features respectively. Another idea is to set the range of each feature to the interval [0,1]. This is called linear normalization which is formulized below.

$$\hat{\mathbf{x}} = \frac{\mathbf{x} - \overline{\overline{x_{min}}}}{\overline{\overline{x_{max}}} - \overline{\overline{x_{min}}}} \quad (3-16)$$

where $\overline{x_{max}}$ and $\overline{x_{min}}$ are the maximum and minimum vectors of the features respectively. Normalizing the feature vectors, instead of normalizing the features, may also be useful. This can simply be performed by setting the magnitude of each feature vector to one. For that reason the feature vectors are divided by their norm as follows:

$$\hat{\mathbf{x}} = \frac{\overline{\mathbf{x}}}{\|\overline{\mathbf{x}}\|} \quad (3-17)$$

3.4 Classification

Classification is assigning class labels to the features extracted from the measurements in the specific problem. This assignment can be performed in a supervised or unsupervised way of learning. In unsupervised learning, any information about the class labels of the measurements is not available even for a small set of data. In supervised learning, there exists a dataset in which the measurements have class labels. In a typical supervised learning procedure, this dataset is divided into two as training set and test set. Using the training set, a classifier is constructed. Then the performance of the classifier is evaluated using the test set. This evaluation is sometimes repeated for different parameters of the classifier constructed. By that way the parameters of the classifier is optimized. After that optimization, the classifier is ready to assign class labels to the features with unknown class labels [57]. Supervised learning is more preferred in the BCI study.

In this thesis, supervised learning algorithms, support vector machine and feed-forward artificial neural network is studied due to their popularity and performance in the SMR-based BCI research. The reader may refer to [58] for a review of classification algorithms used in EEG-based brain computer interfaces.

3.4.1 Support Vector Machines (SVMs)

In this section Support Vector Machines (SVMs) will briefly be explained. To begin with, consider a two class classification problem with 2-dimensional features. Let the circles and triangles in Figure 3-1 represent observations belonging to two different classes. Using these observations, many separating hyperplanes can be selected as classifier for the problem as it is seen in the figure.

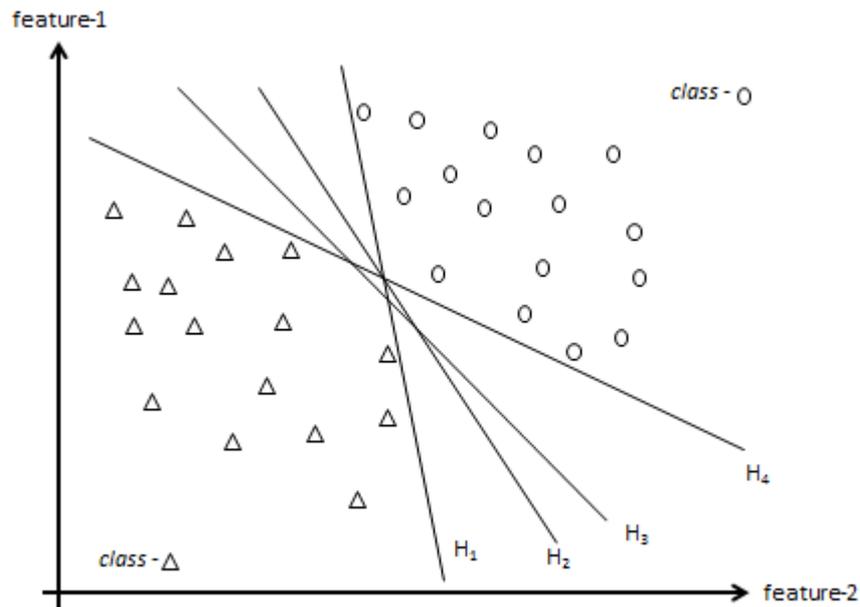


Figure 3-1 : Separating hyperplanes possible to be selected as classifier for the problem

Among all these hyperplanes, SVMs try to find the optimum one which is called Optimum Separating Hyperplane (H_{osh}). H_{osh} is optimum in terms of its generality and robustness. It discriminates the classes such that the margin between the class boundaries is maximized. The class boundaries are determined by the observations closest to H_{osh} which are called support vectors. Support vectors, H_{osh} , and the margin width are shown in Figure 3-2.

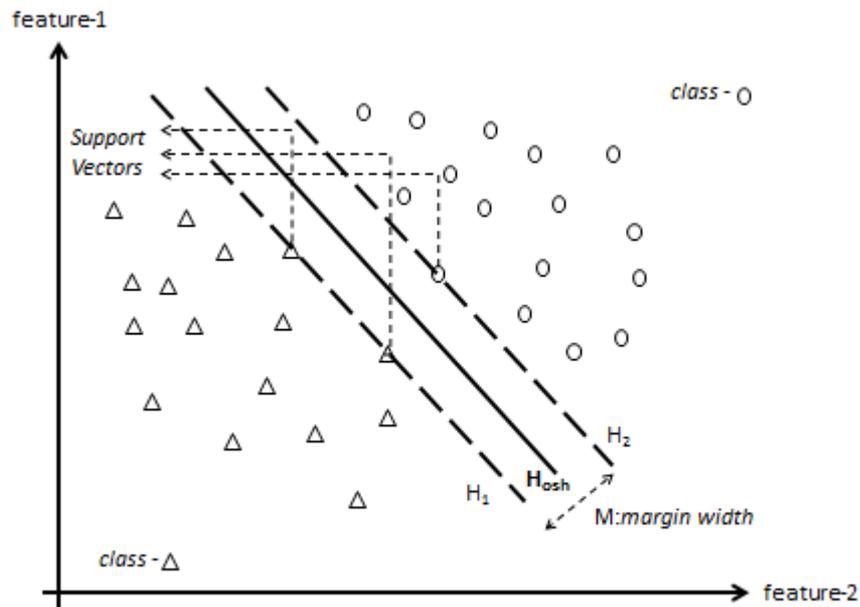


Figure 3-2 : Optimum separating hyperplane maximizes the margin width determined by the support vectors.

The objective function of the SVM algorithm to be minimized can be expressed as follows:

$$\frac{1}{M} + C \sum_{k=1}^R \xi_k \quad (3-18)$$

where

M : The margin width,

ξ_k : The distance of the misclassified observation to its class boundary (see Figure 3-3),

C : Tradeoff parameter between the addends.

The first term in (3-18) is due to maximize the margin width and the second term is to minimize the distance of the misclassified observations to their class boundary. C , the tradeoff parameter between the terms, is selected by hand according to the problem.

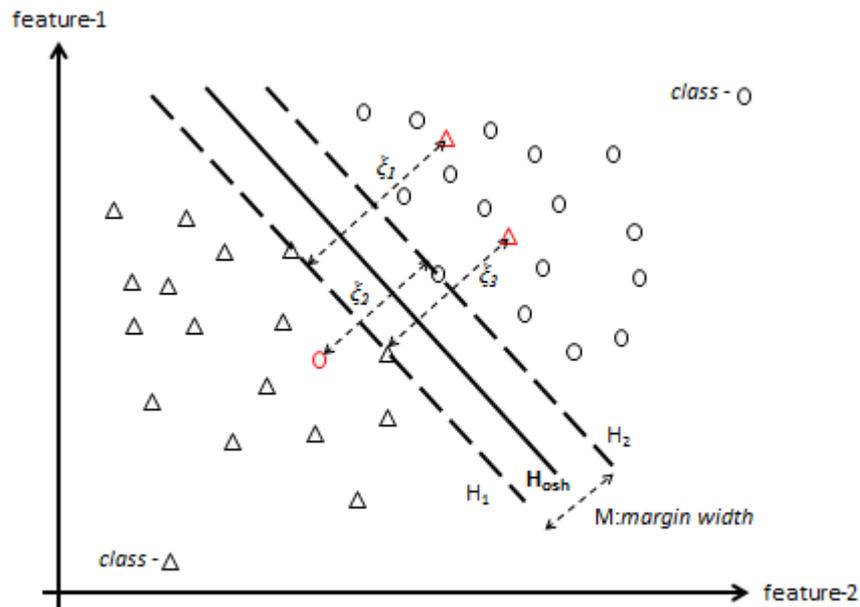


Figure 3-3 : The objective of the SVM algorithm is to maximize the margin width and minimize the distance of the misclassified observations to their class boundary.

When the observations are separable, they are separated in their original space by the H_{osh} . Otherwise, they can be mapped to a higher dimensional space in which they are separable. This situation is illustrated for 1-dimensional feature case below. The observations in Figure 3-4 are separable by a 1D H_{osh} . However it is not the case in 1D for the observations in Figure 3-5. They become separable only when they are mapped to a higher dimensional space as it is seen in Figure 3-6.

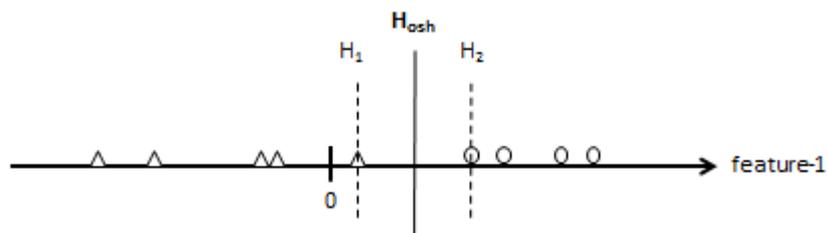


Figure 3-4 : Observations separable in 1D

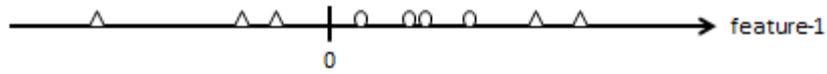


Figure 3-5 : Observations not separable in 1D

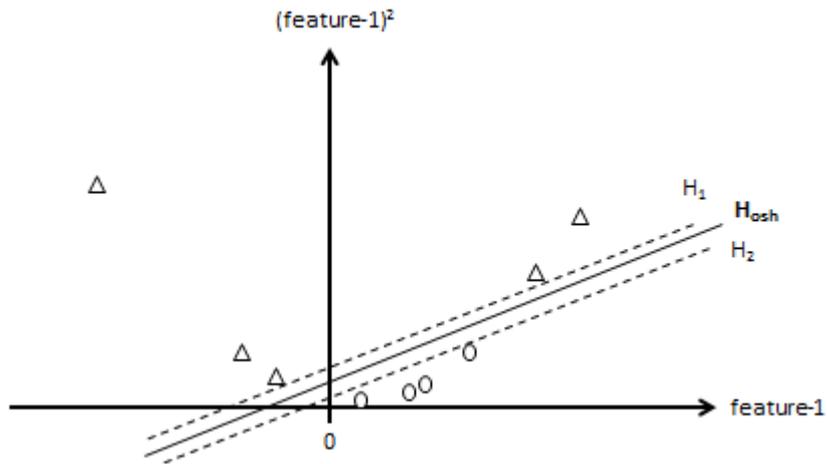


Figure 3-6 : Observations that are not separable in 1D being separable in 2D.

The functions that map the observations to a higher dimensional space are called Kernel functions. Some examples of the Kernel functions used in SVM are given below.

- **Linear Kernel:**

$$K(x_i, x_j) = x_i^T x_j \quad (3-19)$$

- **Polynomial Kernel:**

$$K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \quad \gamma > 0 \quad (3-20)$$

- **Radial Basis Function Kernel:**

$$K(x_i, x_j) = \exp(-\gamma|x_i - x_j|^2), \quad \gamma > 0 \quad (3-21)$$

- **Sigmoid Function:**

$$K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r) \quad (3-22)$$

Here, γ , r , and d are kernel parameters to be adjusted for the specific classification problem [59].

In this thesis study, a well-known SVM toolbox, LIBSVM, is utilized for SVM classification [60].

3.4.2 Artificial Neural Networks (ANNs)

Artificial Neural Networks (ANNs) are one of the non-linear classification algorithms. There are several ANN topologies used in the literature. In this thesis a three-layer feed-forward ANN with one hidden layer and one output layer is implemented and used. The topology of the network is shown in Figure 3-7.

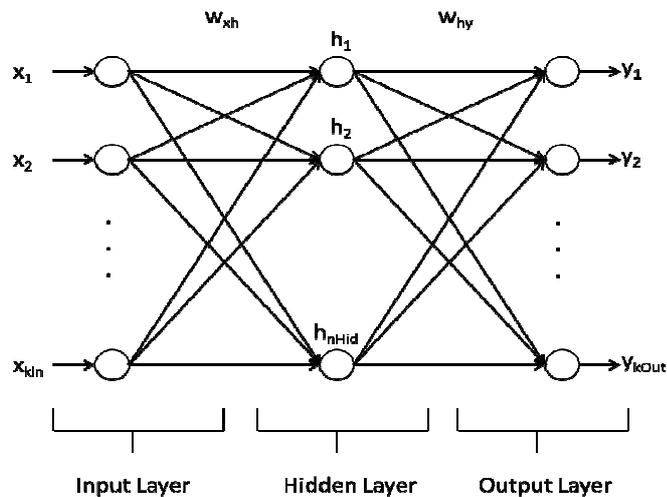


Figure 3-7 : Three layer feed-forward artificial neural network

The terms k_{In} , n_{Hid} , and k_{Out} in Figure 3-7 represent the number of input nodes (i.e. feature dimension), the number of hidden nodes, and the number of output nodes (i.e. number of classes in the problem) respectively. The network has two modes of operation. The first one is the feed-forward mode. It consists of presenting a pattern to the input nodes and passing the signals through the network in order to get output. The second mode of operation is learning which consists of presenting input patterns and finding the network parameters (weights) that minimize the distance between the computed output and the desired output.

3.4.2.1 Feed-forward Operation

In this operation mode, an input pattern is applied to the input layer. Each feature of the pattern is multiplied with a weight and distributed to each unit in the hidden layer. Then, the weighted sum of the features are transformed by a nonlinear activation function. In this study the sigmoid function given in (3-23) is used as the activation function.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3-23)$$

Afterwards, the output of the hidden layer is multiplied with the network weights again and transferred to the output layer. Similar to the hidden layer, a summation and activation operation is performed in the output layer to produce the output vector. The output vector represents the classes in the problem. This representation is performed by unit vectors of which element are 1 only for the related class. For example, for a three class problem, the output vectors representing the classes are [1 0 0], [0 1 0], and [0 0 1].

3.4.2.2 Learning

In order the ANN to be used in feed-forward operation mode, the weights in the network, W_{xh} and W_{hy} , must be calculated. This calculation is performed by back propagation algorithm using the observations in the training data set. Back propagation is a batch training algorithm in which weights are only updated after all the inputs and targets are presented. The difference between the current network output and the desired network output is defined as the error function. The algorithm tries to minimize that function using gradient descent algorithm. The error is a function of network weights. Therefore the derivative of the error function with respect to the network weights is calculated and used in the gradient descent algorithm. The error function to be minimized is given in (3-24).

$$E = \frac{1}{2} (\bar{t} - \bar{o})^2 \quad (3-24)$$

where \bar{t} is the desired output and \bar{o} is the current output. The derivative of the error, E , with respect to the weight from neuron k to j , w_{jk} , and the derivative of the error, E , with respect to the weight from neuron i to k , w_{ki} , are given in (3-25) and (3-26) respectively for the network part given in Figure 3-8.

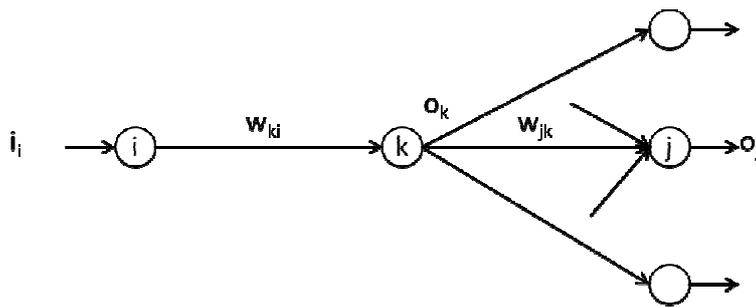


Figure 3-8 : ANN piece

$$\frac{\partial E^p}{\partial w_{jk}} = (o_j^p (1 - o_j^p))(t_j^p - o_j^p) i_k^p \quad (3-25)$$

$$\frac{\partial E^p}{\partial w_{ki}} = (o_k^p(1 - o_k^p)) \left(\sum_{j=1}^c (o_j^p(1 - o_j^p))(t_j^p - o_j^p) w_{jk} \right) i_i^p \quad (3-26)$$

The derivative terms in (3-25) and (3-26) are calculated for each observation in the training set and subtracted from the related weights. The operation continues iteratively until the error term becomes insignificant. The reader may refer to [57] for the derivation and detailed information about the algorithm.

3.5 Evaluation

In order to analyze the performance of BCI systems, several evaluation techniques can be used. In this study classification accuracy, Cohen's Kappa Coefficient, and Nykopp's information transfer are used to analyze the performance in the experiments. They are commonly used in the BCI competitions to compare the results of different research groups [61,62]. The definitions of these evaluation methods are given in the following subsections after defining the terminology in a confusion matrix.

3.5.1 The Confusion Matrix

The confusion matrix consists of elements, n_{ij} , which represent the number of samples of class i predicted as class j . The diagonal elements of the matrix, n_{ii} , show the number of correctly classified samples. The number of samples is calculated as:

$$N = \sum_{i=1}^M \sum_{j=1}^M n_{ij} \quad (3-27)$$

Although confusion matrices give an idea about the performance of the BCI system, they are rarely presented. The summary statistics such as classification accuracy and Cohen's Kappa Coefficient are much more preferred [63].

3.5.2 Classification Accuracy

The classification accuracy (Acc) is the simplest and most widely used way of evaluating a BCI. It is calculated as follows.

$$Acc = \frac{\sum_{i=1}^M n_{ii}}{N} \quad (3-28)$$

On the other hand, there are some limitations of the classification accuracy. First of all it does not consider the off-diagonal elements in the confusion matrix. Also the weight of a class in the calculation depend on the number of samples from that class [63].

3.5.3 Cohen's Kappa Coefficient

When the limitations of the classification accuracy is considered, Cohen's kappa coefficient, κ , serves a more reliable and sensitive evaluation criteria. In the calculation of κ , the classification accuracy, Acc (overall agreement), and the chance agreement, p_e , is used together. The definition of p_e is given below.

$$p_e = \frac{\sum_{i=1}^M n_{.i} n_{i.}}{N^2} \quad (3-29)$$

where $n_{.i}$ and $n_{i.}$ are the sum of the i^{th} column and the i^{th} row of the confusion matrix, respectively. Then, the kappa coefficient, κ , is calculated as it is given below.

$$\kappa = \frac{p_o - p_e}{1 - p_e} \quad (3-30)$$

The maximum value that the kappa coefficient can take is 1 (perfect classification). The value changes depending on the correlation between the predicted classes and the actual classes [63].

3.5.4 Nykopp's Information Transfer

BCIs are alternative communication channels between the brain and the environment. Therefore, the information transfer in that communication channels must be quantified. For that purpose, Nykopp derived information transfer for a general confusion matrix as given below [63][64].

$$I(X; Y) = H(Y) - H(Y|X) \quad (3-31)$$

where $H(x)$ represents the entropy of the discrete random variable x . Specifically;

$$H(Y) = - \sum_{j=1}^M p(y_j) \cdot \log_2 p(y_j) \quad (3-32)$$

with

$$p(y_j) = \sum_{i=1}^M p(x_i) \cdot p(y_j|x_i) \quad (3-33)$$

and

$$H(Y|X) = - \sum_{i=1}^M \sum_{j=1}^M p(x_i) \cdot p(y_j|x_i) \cdot \log_2 p(y_j|x_i) \quad (3-34)$$

In the equations above, the random variable X models the user intention and the random variable Y models the classifier output. M is the number of classes. $p(x_i)$ is the a priori probability for class x_i , $p(y_j)$ is the probability of classifier output to be class y_j , and $p(y_j|x_i)$ is the probability to classify x_i as y_j .

CHAPTER 4

EXPERIMENTS AND RESULTS

4.1 Experiments on BCI Competition III: Dataset V

In this section the experiments performed on the dataset provided by IDIAP Research Institute [65] for BCI Competition III [61] is presented. The dataset contains EEG of three mental tasks which are

1. Imagination of repetitive self-paced left hand movements,
2. Imagination of repetitive self-paced right hand movements,
3. Generation of words beginning with the same random letter.

4.1.1 Explanation of the Experiment

Three normal subjects sitting in a normal chair with relaxed arms resting on their legs attended 4 non-feedback sessions. The first three sessions are provided as *training data* with class labels in order to be used to construct a classifier. The final session is provided as *test data* in order to be used to evaluate the performance of the classifier. The class labels of the test data is announced after the deadline of the competition.

In the experiments, each subject attended 4 sessions separated with 5-10 minutes breaks on the same day. Each session lasted 4 minutes. The subject performed a

task randomly requested by the operator for about 15 seconds then switched to the next task requested by the operator without giving any break.

32 electrodes Biosemi system [66] is used to record the EEG data. The electrodes are located according to the International 10-20 system [67] (see Figure 4-1). The EEG signal is sampled at 512 Hz. Any artifact rejection or correction was not employed on the data.

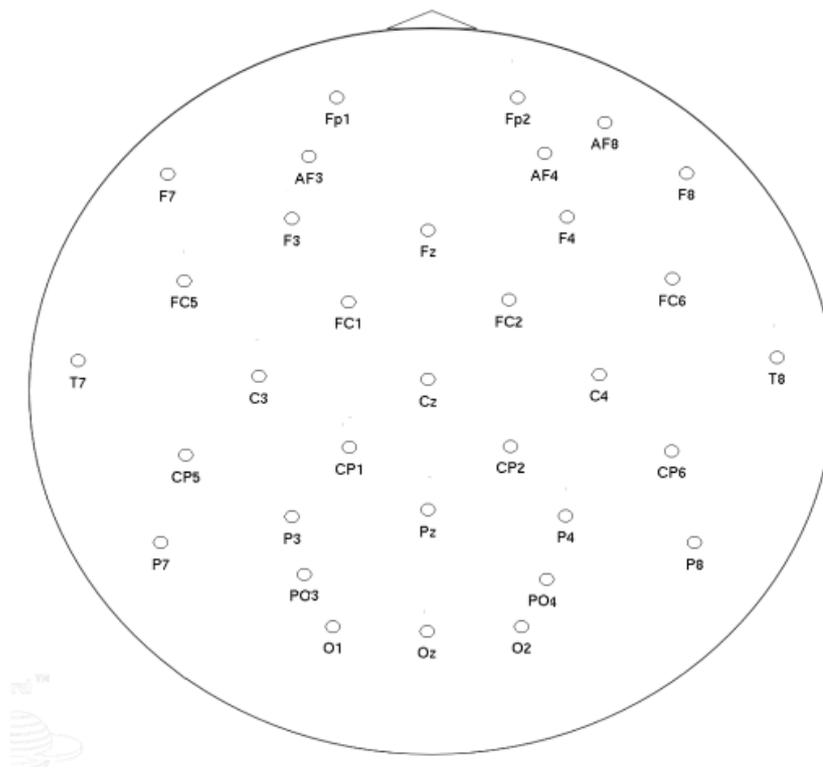


Figure 4-1 : The placement of the electrodes in the experiment

The data for classification are provided in two ways. First one is the precomputed feature vectors and the second one is the raw EEG signals.

4.1.2 Precomputed Features

4.1.2.1 Explanation of the features

Precomputed features are useful in terms of focusing into the classification stage of the EEG signal processing problem without considering the signal enhancement and feature extraction stages. These already extracted features basically summarize the related frequency content of the EEG data. While extracting these features, a surface Laplacian filter [35] was employed on the EEG data first. Then, the power spectral density (PSD) of the filtered EEG data was calculated in the 8-30 Hz band at every 62.5 ms using the last second of the data. The frequency resolution was 2 Hz and the number of electrodes used in PSD calculation was 8 (C3, Cz, C4, CP1, CP2, P3, Pz, P4). As a result, 96 dimensional feature vectors (8 channels times 12 frequency components) were obtained.

4.1.2.2 Results on Precomputed Features

In this part the results obtained using the normalization and classification techniques mentioned in CHAPTER 3 will be presented. For normalization, Linear Feature Normalization (LFN), Gaussian Feature Normalization (GFN), and Unit-norm Feature Vector Normalization (UFVN) are tested. In the classification step, Support Vector Machines (SVM), and Artificial Neural Networks (ANN) are studied. Also Principal Component Analysis (PCA) is utilized in order to reduce the dimension of the precomputed PSD features.

The resultant performance of the methodologies is evaluated on the test data. This strategy enables a comparison of the results with the other methodologies evaluated on the same test data in the scope of the competition. All the parameter optimizations of the methodologies are performed in the training data. For that purpose, randomly selected 75% of the training data is used to construct a classifier and the remaining 25% is used to validate the performance of the model

for the given parameters. While dividing the training data into partitions, uniform distribution of the features for different classes is considered. The final classification model is constructed with the parameters giving maximum validation result using whole training data. Then, the performance of the model is evaluated on the test data.

In the evaluation step, a classification output for each input feature vector, computed 16 times per second, is generated first. Then the average of 8 consecutive outputs is calculated in order to produce a response at every 0.5 seconds. Finally, classification accuracy is calculated for these responses as mentioned in the requirements of the competition.

Validation Results for SVM Classification

In SVM classification, the RBF kernel is preferred since it is reported to provide the best results in terms of the classification performance [68]. For the PCA based RBF kernel SVM classification, there are three parameters to be optimized. First one is the γ coefficient of the RBF kernel, second one is the regularization parameter C, and the third one is the PCA-Coefficient. Since there is no direct analytical way of finding the optimum values of these parameters, *validation accuracy* is calculated in the training data for different γ , C, and PCA-Coefficient combinations. The combination giving the maximum validation accuracy is used to construct the final classification model. In this study γ and C values are grown exponentially in order to enlarge the search space. γ is varied between 2^{-31} and 2^5 while C is varied between 2^{-1} and 2^7 . An example validation accuracy table is given in Table 4-1. In the table, validation accuracies for different subjects under the same normalization and PCA procedures are given.

Table 4-1 : Validation accuracies for all subjects and different γ & C combinations. The PCA-Coefficient is 99 and unit norm feature vector normalization and no feature normalization is used.

		log(γ):	-31	-27	-23	-19	-15	-11	-7	-3	1	5
subject	log(C)											
1	-1	39,36	39,36	39,36	39,36	39,36	39,36	39,36	70,33	74,05	97,23	39,97
	3	39,36	39,36	39,36	39,36	39,36	39,36	70,21	72,23	84,08	99,58	50,80
	7	39,36	39,36	39,36	39,36	70,21	72,07	74,09	93,09	99,58	50,80	
2	-1	38,77	38,77	38,77	38,77	38,77	38,77	38,77	56,96	65,35	98,23	39,19
	3	38,77	38,77	38,77	38,77	38,77	57,15	63,31	81,62	99,69	43,62	
	7	38,77	38,77	38,77	38,77	57,19	63,19	66,31	92,54	99,69	43,62	
3	-1	49,03	49,03	46,42	46,62	46,66	46,66	54,00	58,94	98,56	33,86	
	3	49,03	49,03	46,42	46,62	46,66	54,16	55,95	75,04	99,81	73,56	
	7	49,03	49,03	46,42	46,62	54,16	55,83	58,71	89,66	99,81	73,56	

It is seen that C value does not affect the validation accuracy as much as the γ value. It is also seen that the optimum γ &C combination does not depend on the subject. However, it is seen in Table 4-2 that the optimum γ &C combination depend on the feature vector normalization (FVN) and feature normalization (FN) methods. In the table, the validation accuracies for different normalization methods and γ &C combinations for subject 1 with PCA-Coefficient 99.

Table 4-2 : The validation accuracies for different normalization methods and γ &C combinations for subject 1 with PCA coefficient 99.

		log(γ):	-31	-27	-23	-19	-15	-11	-7	-3	1	5
FVN-FN	log(C)											
NoFVN-GFN	-1	39,36	39,36	39,36	39,36	39,36	39,36	55,40	76,37	98,10	39,36	39,36
	3	39,36	39,36	39,36	39,36	55,55	72,45	88,72	99,62	41,98	39,36	
	7	39,36	39,36	39,36	55,51	71,85	77,51	94,98	99,62	41,98	39,36	
NoFVN-LFN	-1	39,36	39,36	39,36	39,36	39,36	39,36	39,36	53,84	72,95	92,90	42,52
	3	39,36	39,36	39,36	39,36	39,36	54,45	72,19	79,64	99,09	71,77	
	7	39,36	39,36	39,36	39,36	54,56	72,07	72,95	89,78	99,13	71,77	
NoFVN-NoFN	-1	39,36	39,36	39,36	39,36	56,61	72,87	85,30	87,84	39,36	39,36	
	3	39,36	39,36	39,36	56,69	72,72	74,89	97,68	99,43	39,40	39,36	
	7	39,36	39,36	56,69	72,68	72,99	85,49	98,63	99,43	39,40	39,36	
UFVN-GFN	-1	39,36	39,36	39,36	39,36	70,02	92,02	42,78	39,36	39,36	39,36	
	3	39,36	39,36	39,36	69,79	78,42	99,16	76,33	39,36	39,36	39,36	
	7	39,36	39,36	69,87	72,64	88,68	99,16	76,33	39,36	39,36	39,36	
UFVN-LFN	-1	39,36	39,36	39,36	39,36	39,36	39,36	69,34	72,87	96,47	40,20	
	3	39,36	39,36	39,36	39,36	39,36	69,45	71,77	82,22	99,54	55,32	
	7	39,36	39,36	39,36	39,36	69,45	71,81	72,80	91,83	99,58	55,32	
UFVN-NoFN	-1	39,36	39,36	39,36	39,36	39,36	39,36	70,33	74,05	97,23	39,97	
	3	39,36	39,36	39,36	39,36	39,36	70,21	72,23	84,08	99,58	50,80	
	7	39,36	39,36	39,36	39,36	70,21	72,07	74,09	93,09	99,58	50,80	

In Table 4-3 the validation accuracies for different PCA-Coefficients and γ &C combinations for subject 1 is given. For the results in the table, the feature normalization type is Gaussian and no feature vector normalization is employed. It is seen in the table that the optimum γ &C combination strongly depend on the PCA-Coefficient. This situation is the reason of searching optimum γ &C combination in a range that wide.

Table 4-3 : The validation accuracies for different PCA-Coefficients and γ &C combinations for subject 1. The feature normalization type is Gaussian and no feature vector normalization is employed.

	log(γ):	-31	-27	-23	-19	-15	-11	-7	-3	1	5
PCA-Coeff.	log(C)										
97	-1	39,36	39,36	39,36	39,36	39,36	39,36	62,20	64,44	64,44	64,36
	3	39,36	39,36	39,36	39,36	39,36	62,35	64,70	64,89	64,86	64,13
	7	39,36	39,36	39,36	39,36	62,39	64,78	64,51	64,29	64,93	64,10
97,5	-1	39,36	39,36	39,36	39,36	39,36	41,19	67,21	68,88	69,87	70,10
	3	39,36	39,36	39,36	39,36	41,19	67,36	67,71	69,22	70,63	66,49
	7	39,36	39,36	39,36	41,19	67,33	67,44	68,96	69,22	70,86	64,36
98	-1	39,36	39,36	39,36	39,36	39,36	48,48	68,35	72,15	82,26	39,36
	3	39,36	39,36	39,36	39,36	48,56	68,31	70,44	75,95	87,92	42,06
	7	39,36	39,36	39,36	48,56	68,35	69,15	71,73	78,69	87,65	42,06
98,5	-1	39,36	39,36	39,36	39,36	39,36	49,43	70,59	83,85	46,16	39,36
	3	39,36	39,36	39,36	39,36	49,51	69,83	74,39	93,35	83,78	39,40
	7	39,36	39,36	39,36	49,51	69,60	71,09	79,71	93,66	83,78	39,40
99	-1	39,36	39,36	39,36	39,36	39,36	55,40	76,37	98,10	39,36	39,36
	3	39,36	39,36	39,36	39,36	55,55	72,45	88,72	99,62	41,98	39,36
	7	39,36	39,36	39,36	55,51	71,85	77,51	94,98	99,62	41,98	39,36
99,5	-1	39,36	39,36	39,36	39,36	39,36	63,41	86,36	53,53	39,36	39,36
	3	39,36	39,36	39,36	39,36	62,46	76,90	97,68	98,06	39,36	39,36
	7	39,36	39,36	39,36	62,16	74,32	87,39	98,18	98,06	39,36	39,36
100	-1	49,85	66,95	94,49	39,48	39,36	39,36	39,36	39,36	39,36	39,36
	3	65,20	86,47	99,58	46,66	39,36	39,36	39,36	39,36	39,36	39,36
	7	77,81	96,69	99,58	46,66	39,36	39,36	39,36	39,36	39,36	39,36

Another property observed in Table 4-3 is the general tendency to increase in the maximum validation accuracies which are marked with boxes. This is expected since the feature vector dimension and the percent of information represented by feature vectors increases while the PCA-Coefficient is increasing. However, the maximum validation accuracy is obtained for the PCA-Coefficient 99. The reason is the fact that some part of the information lost in PCA is noise. This is the main

advantage of using PCA in classification problems. For smaller PCA-Coefficients, the percentage of noise in the lost information decreases. Therefore a decrease in the validation accuracy is observed. However, PCA may still be preferred in the cases where high dimensionality is a problem in terms of generating fast responses.

In Table 4-4, the feature vector dimensions for different PCA-Coefficients are given for each subject. The PCA coefficient is a measure of percent information maintained after the PCA operation. For example, the 97% percent of the information in 96-dimensional feature space can be represented in 2-dimensional feature space for the subject 1. The situation is similar for the other subjects. This shows the high correlation between the features. This correlation arises from the blurring of the EEG signal in the skull and low spatial resolution of the signal acquisition methodology.

Table 4-4 : Feature vector sizes for different PCA-Coefficients.

subject/PCA-coefficient	97	97,5	98	98,5	99	99,5	100
1	2	4	9	17	36	59	96
2	4	6	12	23	39	61	96
3	3	6	12	25	39	63	96

Validation Results for ANN Classification

For ANN classification, there are two parameters to be optimized. First one is the number of nodes in the hidden layer, nHidden, of the three layer feed-forward network and the second one is the PCA-Coefficient. For that purpose, validation accuracy is calculated in the training data for different nHidden and PCA-Coefficient combinations. The combination giving the maximum validation accuracy is used to construct the final classification model. In this study nHidden is varied between 5 and 33 with step size 4. In Table 4-5, validation accuracies for different subjects and nHidden & PCA-Coefficient combinations are given. For

the results in the table, linear feature normalization and unit norm feature vector normalization is used.

Table 4-5 : The validation accuracies for different subjects and nHidden & PCA-Coefficient combinations. Linear feature normalization and unit norm feature vector normalization is used.

	PCA-Coefficient:	97	97,5	98	98,5	99	99,5
1	nHidden						
	5	61,97	66,11	62,99	68,47	70,55	73,59
	9	61,70	66,19	65,81	68,35	69,30	75,11
	13	61,44	66,45	65,01	68,77	70,06	74,13
	17	61,74	66,38	65,65	68,69	70,52	75,15
	21	61,63	66,45	66,87	69,26	70,48	74,58
	25	61,85	66,95	66,83	69,76	70,02	74,77
	29	61,51	66,38	66,26	68,92	70,10	75,27
	33	61,85	66,64	66,68	68,85	70,29	75,76
2	5	54,27	54,12	54,15	55,58	61,54	63,85
	9	54,65	54,77	54,58	59,00	62,65	64,08
	13	54,73	54,08	54,38	57,96	62,04	63,15
	17	54,77	54,27	55,38	57,58	62,27	63,73
	21	54,65	54,15	55,12	58,54	62,65	63,81
	25	54,42	54,38	55,54	58,35	62,73	64,31
	29	55,00	54,69	55,42	58,73	62,00	64,73
	33	54,92	54,46	55,69	58,38	62,58	64,46
3	5	44,79	45,92	46,54	51,24	47,47	52,18
	9	45,88	46,31	46,85	51,44	52,72	55,25
	13	45,72	46,73	46,31	51,87	53,34	55,37
	17	45,65	46,50	47,12	51,83	55,40	55,68
	21	46,11	46,70	46,81	51,67	54,16	56,18
	25	45,10	46,19	47,67	52,33	53,46	55,64
	29	45,10	46,93	47,78	52,33	54,20	56,69
	33	45,61	46,46	47,55	53,15	54,74	56,10

It is seen in Table 4-5 that there is a general tendency to increase in the validation values with the increasing nHidden and PCA-Coefficient values. However, it is seen that the validation accuracy value does not change that much for nHidden

values greater than 20 under the same PCA procedure. 21 hidden nodes seem to be enough to perform the required classification operation in the problem. Any nHidden value greater than 20 may give the maximum validation accuracy as it is marked with boxes in the table.

The training of an ANN with back propagation algorithm is an iterative procedure as it is mentioned in section 3.4.2. In these iterations, the validation accuracy must be checked against memorization of the neural network. Memorization occurs when the classification model is too complex to classify all the data in the training set with losing its generality. Due to this loss in generality, the network becomes useless for any other data except from the training set. Validation accuracy gives an idea about the generality of the model since it is calculated using data different from the data training the ANN. In Figure 4-2, the root-mean-square error (RMSE) calculated using 75% of the training data; in Figure 4-3, validation accuracy calculated using the remaining 25% at each iteration is given for subject 1. The formulation of RMSE is given in (4-1).

$$\text{RMSE} = \frac{1}{nF} \sum_{i=1}^{nF} \sum_{j=1}^{nOut} (t_{ij} - y_{ij})^2 \quad (4-1)$$

where nF is the number of features, nOut is the number of output nodes (which is 3 since there are 3 output classes in this problem), t_{ij} is the current output of the network and y_{ij} is the desired output of the network.

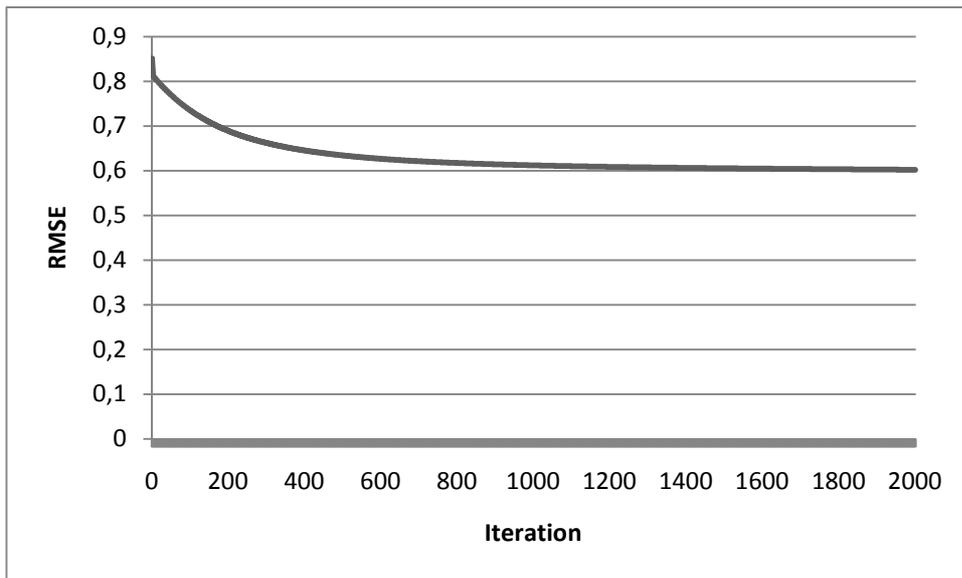


Figure 4-2 : RMS Error calculated at each iteration for subject 1 under linear feature normalization and unit norm feature vector normalization with PCA-Coefficient 99.5.

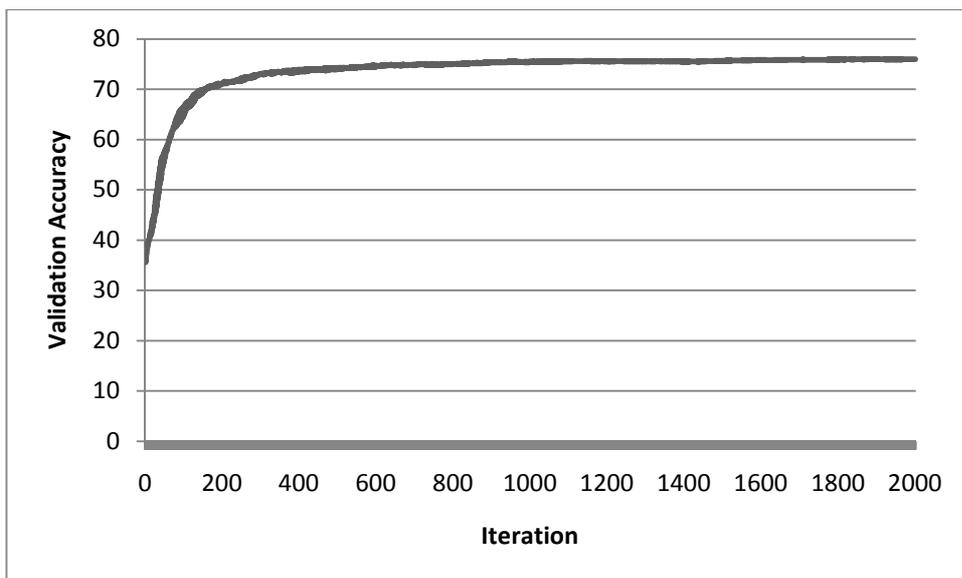


Figure 4-3 : Validation accuracy calculated at each iteration for subject 1 under linear feature normalization and unit norm feature vector normalization with PCA-Coefficient 99.5.

The non-decreasing characteristic of the validation accuracy curve in Figure 4-3 shows that there is no memorization problem while the RMSE of the training set is decreasing. This may be due to the large number of feature vectors in the

training set. As a result, constant and sufficient number of iterations can be used as stopping criteria in the training procedure. 1500 iterations is used in the experiments which seem to be sufficient in the Figure 4-3.

Overall Evaluation Results

After optimizing the parameters using validation accuracies in the training set; the normalization, feature extraction and classification methodology is evaluated in the test set. The number of NN and SVM calculations in this optimization process are given in Table 4-6.

Table 4-6 : The number of parameters and total number of combinations.

	nHidden	γ	C	FVN	FN	PCA	Total # of combinations
NN	8	-	-	2	3	7	336
SVM	-	10	3	2	3	7	1260

With the parameters giving the maximum validation accuracy, each methodology is evaluated. In Table 4-7, the classification accuracies calculated using the responses computed 16 times per second (16 Hz accuracy results); in Table 4-8, the classification accuracies calculated using the responses computed 2 times per second (2 Hz accuracy results) from the average of 8 consecutive outputs are given.

Table 4-7 : Classification accuracy calculated using the responses computed 16 times per second

	FN:	GFN			LFN			NO		
	PCA:	YES		NO	YES		NO	YES		NO
	FVN:	NO	UFVN	NO	NO	UFVN	NO	NO	UFVN	NO
classifier	subject									
NN	1	72,29	71,26	65,04	74,14	74,34	74,14	72,15	73,74	69,86
NN	2	57,40	56,19	51,73	60,54	59,53	60,54	57,03	58,99	53,92
NN	3	49,40	47,91	42,57	50,72	52,32	50,72	48,45	52,67	47,39
NN	Avg	59,70	58,45	53,11	61,80	62,07	61,80	59,21	61,80	57,06
SVM	1	69,75	71,18	64,33	73,57	74,34	73,57	72,03	75,14	72,03
SVM	2	53,60	54,41	51,70	59,76	56,45	59,76	57,89	57,57	57,89
SVM	3	46,39	46,96	46,39	48,68	49,97	48,68	47,33	49,68	47,33
SVM	Avg	56,58	57,51	54,14	60,67	60,26	60,67	59,09	60,80	59,09

Table 4-8 : Classification accuracy calculated using the responses computed 2 times per second from the average of 8 consecutive outputs

	FN:	GFN			LFN			NO		
	PCA:	YES		NO	YES		NO	YES		NO
	FVN:	NO	UFVN	NO	NO	UFVN	NO	NO	UFVN	NO
classifier	subject									
NN	1	76,03	73,74	70,09	76,71	77,40	75,34	76,48	76,26	71,46
NN	2	60,37	61,75	54,15	61,75	60,83	61,75	60,83	61,06	55,07
NN	3	52,98	52,75	41,51	53,67	55,50	52,52	51,61	55,28	48,62
NN	Avg	63,13	62,75	55,25	64,04	64,58	63,21	62,97	64,20	58,38
SVM	1	73,52	73,97	67,58	75,57	78,54	75,57	75,11	79,68	75,11
SVM	2	55,76	56,68	53,23	62,90	59,91	62,90	58,76	61,06	58,76
SVM	3	50,92	49,08	50,92	51,83	53,90	51,83	51,61	51,83	51,61
SVM	Avg	60,06	59,91	57,24	63,44	64,12	63,44	61,83	64,19	61,83

It is obvious in the resulting tables that the averaging operation increases the classification accuracy. However, the 16 Hz accuracy results are more sensitive and useful while comparing the methodologies.

Maximum evaluation accuracy is obtained with NN employing PCA in the feature extraction, with linear feature normalization (LFN) and unit-norm feature vector normalization (UFVN). PCA and UFVN give good results also in SVM classification.

4.1.3 Raw EEG Signals

In addition to the precomputed features, the 32 channel EEG recordings, that the PSD features are extracted from, are also provided in the competition. These signals enable to test different preprocessing and feature extraction techniques.

4.1.3.1 Explanation of the features

In the experiments on raw EEG signals, CSP features are extracted using different spatial filtering, temporal filtering and normalization techniques. Then the results are compared with the results on precomputed PSD features.

Similar to the PSD feature extraction strategy, CSP features are calculated at every 62.5 ms (i.e., 16 times per second) over the last second of data using the same EEG channels (C3, Cz, C4, CP1, CP2, P3, Pz, and P4). The other electrode recordings provided are only used in the spatial filtering step.

4.1.3.2 Results on Raw EEG signals

In this section the results obtained using different spatial filtering, temporal filtering, feature normalization, feature vector normalization and classification techniques will be presented. The list of the techniques utilized with different combinations is given in Table 4-9.

Table 4-9 : The signal processing techniques used to classify the CSP features extracted from the raw EEG signal

Spatial filter (SF):	Common Average Reference (CAR) filter Large Laplacian (LL) filter Small Laplacian (SL) filter no (NO) filter
Temporal filter (TF):	8-30 Hz bandpass filter (YES) no (NO) filtering
Feature Normalization (FN):	Linear Feature Normalization (LFN) Gaussian Feature Normalization (GFN) No (NO) Normalization
Feature Vector Normalization (FVN):	Unit-norm Feature Vector Normalization (UFVN)
Classification:	SVM, ANN

The validation and the evaluation procedure is same with the procedure followed for the precomputed PSD features.

For CSP based SVM classification, there are three parameters to be optimized. First one is the γ coefficient of the RBF kernel, second one is the regularization parameter C , and the third one is p which is the number of eigenvectors used to construct the CSP filter W . In order to optimize these parameters, randomly selected 75% of the training data is used to construct a classifier and the remaining 25% is used to validate the performance of the model for the given the $\gamma - C - p$ combination. While dividing the training data into partitions, uniform distribution of the features for different classes is considered.

For CSP based NN classification, there are two parameters to be optimized. First one is the number of hidden nodes, n_{Hidden} , and the second one is p . These parameters are optimized in the same way with the CSP based SVM classification.

The final classification model is constructed with the parameters giving maximum validation accuracy value using whole training data. Then, the performance of the model is evaluated on the evaluation data. The evaluation results with SVM for different filtering and normalization techniques are given in Table 4-10 and the evaluation results with ANN for different filtering and normalization techniques are given in Table 4-11.

Table 4-10 : Evaluation results with SVM for different filtering and normalization techniques

SF	FN	TF: Subject:	YES	YES	YES	YES	NO	NO	NO	NO
			1	2	3	Avg.	1	2	3	Avg.
		FVN								
CAR	GFN	NO	77,92	61,89	51,15	63,66	56,26	54,32	43,19	51,26
CAR	GFN	UFVN	78,77	58,95	50,94	62,89	54,56	49,68	41,51	48,59
CAR	LFN	NO	78,13	63,79	55,14	65,69	57,32	55,37	42,35	51,68
CAR	LFN	UFVN	77,92	63,58	54,93	65,47	54,14	50,32	39,20	47,89
CAR	NO	NO	78,34	63,16	52,20	64,57	56,69	53,26	41,93	50,63
CAR	NO	UFVN	76,01	58,74	52,20	62,32	52,87	51,79	38,16	47,60
LL	GFN	NO	60,30	50,95	48,85	53,36	40,76	45,47	40,04	42,09
LL	GFN	UFVN	66,24	50,95	47,80	55,00	41,19	44,84	41,93	42,65
LL	LFN	NO	63,06	49,68	48,43	53,72	40,55	43,79	39,41	41,25
LL	LFN	UFVN	62,00	51,58	48,22	53,93	43,74	42,11	42,77	42,87
LL	NO	NO	65,39	52,21	47,38	54,99	44,37	44,21	40,67	43,09
LL	NO	UFVN	66,88	52,84	45,28	55,00	43,31	44,63	44,03	43,99
NO	GFN	NO	67,94	56,84	49,06	57,95	40,55	45,68	38,36	41,53
NO	GFN	UFVN	66,03	56,42	46,75	56,40	48,20	47,58	41,09	45,62
NO	LFN	NO	72,40	56,21	49,48	59,36	42,89	46,74	39,62	43,08
NO	LFN	UFVN	71,76	56,42	50,94	59,71	47,56	46,53	41,30	45,13
NO	NO	NO	71,34	56,84	47,38	58,52	44,59	46,32	39,62	43,51
NO	NO	UFVN	70,91	57,89	49,48	59,43	48,41	45,68	40,25	44,78
SL	GFN	NO	58,81	54,95	49,48	54,41	40,13	50,32	39,20	43,22
SL	GFN	UFVN	61,78	55,58	46,96	54,77	45,86	46,74	43,40	45,33
SL	LFN	NO	61,36	54,53	50,10	55,33	40,55	46,74	43,40	43,56
SL	LFN	UFVN	62,85	54,74	50,73	56,11	48,41	46,11	44,03	46,18
SL	NO	NO	66,67	53,26	48,64	56,19	40,34	47,16	46,75	44,75
SL	NO	UFVN	66,03	54,53	49,48	56,68	45,86	45,89	36,27	42,67

Table 4-11 : Evaluation results with ANN for different filtering and normalization techniques

		TF:	YES	YES	YES	YES	NO	NO	NO	NO
		Subject:	1	2	3	Avg.	1	2	3	Avg.
SF	FN	FVN								
CAR	GFN	NO	81,10	66,32	58,07	68,50	60,30	55,79	48,01	54,70
CAR	GFN	UFVN	80,68	65,47	60,17	68,77	56,69	54,95	52,20	54,61
CAR	LFN	NO	81,10	63,37	59,96	68,14	64,33	54,11	50,73	56,39
CAR	LFN	UFVN	81,95	65,68	59,33	68,99	54,14	54,53	47,59	52,09
CAR	NO	NO	83,44	69,68	59,75	70,96	50,96	47,58	48,64	49,06
CAR	NO	UFVN	81,10	63,79	57,65	67,52	48,62	45,68	48,01	47,44
LL	GFN	NO	69,64	54,95	48,43	57,67	45,44	47,58	48,22	47,08
LL	GFN	UFVN	68,79	53,05	48,43	56,76	49,26	45,05	44,86	46,39
LL	LFN	NO	70,91	55,79	50,10	58,94	46,07	48,00	49,69	47,92
LL	LFN	UFVN	71,97	55,16	49,90	59,01	46,28	44,42	47,17	45,96
LL	NO	NO	71,76	53,05	51,36	58,73	48,41	41,26	47,38	45,68
LL	NO	UFVN	71,13	57,05	50,73	59,64	44,80	48,84	44,03	45,89
NO	GFN	NO	78,56	61,05	51,36	63,66	51,80	50,32	45,91	49,34
NO	GFN	UFVN	77,28	60,42	51,15	62,95	50,32	49,89	46,54	48,92
NO	LFN	NO	81,53	59,16	50,52	63,74	56,26	52,42	44,03	50,90
NO	LFN	UFVN	80,04	59,37	51,15	63,52	47,77	49,47	42,77	46,67
NO	NO	NO	80,47	62,53	53,04	65,34	49,68	50,53	42,77	47,66
NO	NO	UFVN	80,89	59,37	51,57	63,94	40,34	47,58	45,70	44,54
SL	GFN	NO	69,64	58,11	53,46	60,40	48,41	55,79	45,28	49,83
SL	GFN	UFVN	68,79	59,58	53,88	60,75	49,47	53,05	44,65	49,06
SL	LFN	NO	74,10	59,16	51,57	61,61	46,07	54,95	41,51	47,51
SL	LFN	UFVN	73,89	60,00	52,83	62,24	47,77	51,79	43,61	47,72
SL	NO	NO	73,25	60,00	54,51	62,59	44,16	53,26	40,04	45,82
SL	NO	UFVN	74,52	60,63	50,52	61,89	40,34	53,89	41,30	45,18

It is obvious in the evaluation results that the temporal filtering increases the classification accuracy as expected. Among the spatial filtering techniques, CAR, also increases the accuracy both for SVM and ANN classifiers. However, the normalization techniques does not affect the accuracy as much as the filtering techniques. The best results are obtained using band-pass filter and CAR filter for both SVM and ANN classifiers.

4.1.4 Conclusion

The maximum accuracy values obtained in this experiment are given with the summary of the methods attended the competition in Table 4-12.

Table 4-12 : The results of BCI competition III

group	Avg.	Subject			Method
		1	2	3	
METU BCI	70.96	83.44	69.68	59.75	CSP+NN
1	68.65	79.60	70.31	56.02	PSD + Distance Based Discriminator
2	68.50	78.08	71.66	55.73	PSD + Feature selection, SVM
Cheng & Ming [45]	68.35	78.31	70.27	56.46	PSD + PCA + improved particle swarm optimization-NN
3	65.90	77.85	66.36	53.44	PSD + Radial Basis Network, SVM
METU BCI	65.69	78.13	63.79	55.14	CSP+SVM
4	65.67	76.03	69.36	51.61	PSD + Fisher's Discriminant Analysis
5	64.91	78.08	63.83	52.75	PSD + Regularized Discriminant Analysis
6	64.60	81.05	73.04	39.68	PSD + Minimum Mahalanobis Distance
METU BCI	64.58	77.40	60.83	55.50	PCA+PSD+NN
METU BCI	64.19	79.68	61.06	51.83	PCA+PSD+SVM
7	64.04	76.06	64.83	51.18	PSD + SVM, CART Decision Tree, LVQ, Naive Bayes
8	63.91	77.40	63.83	50.46	PSD + NN, Linear Discriminant Analysis

CSP features give better results for both SVM and ANN classifications. The best result is obtained with classification of CSP features with ANN. However the SVM results for CSP features are also comparable with the other results in the competition. Furthermore, the training and testing duration of SVM is observed to be 4-5 times shorter than ANN. This makes SVM preferable in online applications considering short response time.

4.2 Experiments on BCI Competition IV: Dataset IIa

In this section the experiments performed on the dataset provided by Graz University of Technology for BCI Competition IV [62] is presented. The dataset contains EEG of four different mental tasks which are

1. Imagination of left hand movements,
2. Imagination of right hand movements,
3. Imagination of feet movements,
4. Imagination of tongue movements.

4.2.1 Explanation of the Experiment

9 subjects sitting in a comfortable armchair attended two sessions on different days. Each session consists of 6 runs separated by short breaks. In each run, the mental tasks were performed 12 times in random order. Therefore, one session contains 288 trials (6 runs x 4 tasks x 12 repetitions). The timing scheme of a single trial is explained below and illustrated in Figure 4-4 [69].

t = 0 s : The subject is warned with an acoustic sound

t = 0-2 s : A fixation cross is presented to the subject

t = 2-3.25 s : A cue related to the task to be performed is presented.

t = 3.25-6 s : The fixation cross is presented.

t = 6-7.5 s : Break before the next trial. The screen is black.

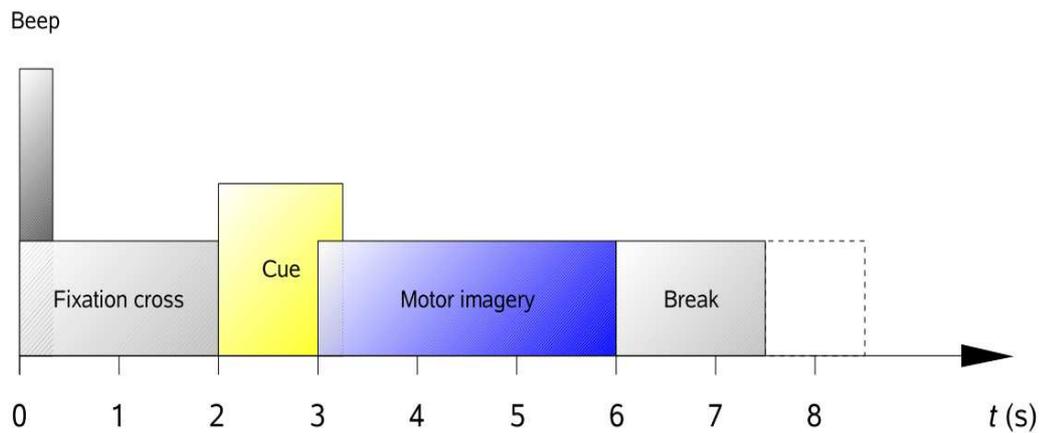


Figure 4-4 : The timing scheme for a single trial of the experiment [69].

The EEG data was recorded using 22 electrodes located at the positions shown in Figure 4-5. The sampling rate was 250 Hz and the EEG signal was band-pass filtered at 0.5-100 Hz. Also a 50 Hz notch filter was used to suppress the line noise [69].

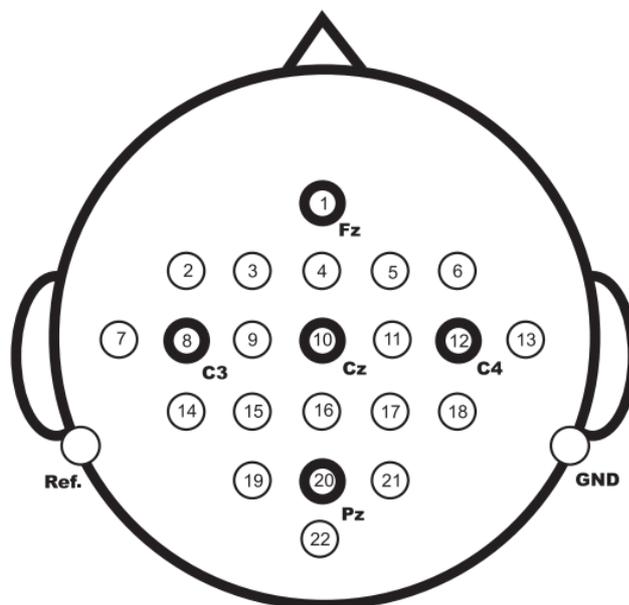


Figure 4-5 : Electrode configuration used in the experiment [69].

4.2.2 Explanation of the Data

Among the provided two sessions the first session is used as training data and the second one is used as evaluation data. The time segments used to extract CSP features for training and evaluation data is shown in Figure 4-6.

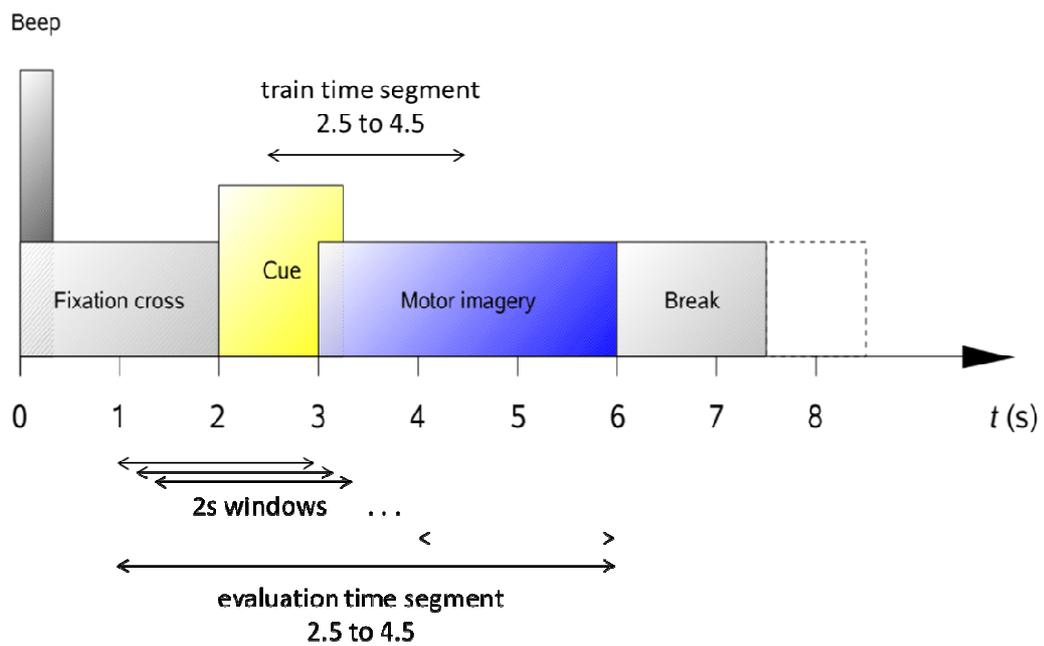


Figure 4-6 : The time segments used in feature extraction for training and evaluation data (adapted from [69]).

CSP features extracted from the 2s window between $t=2.5s$ and $t=4.5s$ in each trial of the *training data* are used to construct SVM classification model. The evaluation features are extracted using the sliding 2s windows between $t=1s$ and $t=6s$ in each trial of the *evaluation data*. Therefore, the classification model produces an output between $t=3s$ and $t=6s$ with a delay of 2s. The windows slide by 10 samples in the *evaluation data*. By this way, an output is generated 25 times per second. This timing scheme is also preferred by the winner algorithm of the competition [70].

4.2.2.1 Results

Validation Results

The validation and the evaluation procedure is same with the procedure followed in the experiments on BCI Competition III dataset V (section 4.1). The sets of parameters used in the validation process are given below.

$$\begin{aligned}\gamma &: \{2^{-10}, 2^{-8}, \dots, 2^2\} \\ C &: \{2^{-1}, 2^3, 2^7\} \\ p &: \{2, 4, \dots, 10\}\end{aligned}\tag{4-2}$$

An example validation accuracy table is given in Table 4-13. In the table, validation kappa values for different γ - C - p combinations are given for subject A09 with CAR spatial filtering, 6-8Hz band-pass filtering, linear feature normalization and no feature vector normalization. As it is seen in the table, there exists 3 different maximum values for the validation kappa value. The values are marked with boxes. In such cases, the combination with the minimum parameter values is selected considering the computational performance in the classification stage.

Table 4-13 : Validation kappa values for subject A01 with different γ - C - p combinations. (spatial filtering: CAR, temporal filtering: 6-8Hz band-pass, feature normalization: linear, feature vector normalization: No)

	log(γ):	-10	-8	-6	-4	-2	0	2
	log(C)							
2	-1	0,63	0,63	0,74	0,80	0,83	0,81	0,81
	3	0,74	0,80	0,83	0,85	0,87	0,85	0,80
	7	0,83	0,85	0,87	0,81	0,81	0,83	0,80
4	-1	0,54	0,54	0,63	0,70	0,76	0,78	0,39
	3	0,63	0,72	0,81	0,81	0,81	0,83	0,80
	7	0,81	0,80	0,81	0,74	0,78	0,83	0,80
6	-1	0,65	0,65	0,72	0,78	0,80	0,80	0,54
	3	0,72	0,80	0,85	0,81	0,81	0,81	0,65
	7	0,87	0,81	0,83	0,83	0,80	0,81	0,65
8	-1	0,50	0,52	0,63	0,67	0,74	0,78	0,19
	3	0,63	0,72	0,83	0,74	0,81	0,80	0,57
	7	0,83	0,74	0,70	0,70	0,81	0,80	0,57
10	-1	0,61	0,65	0,74	0,78	0,78	0,70	0,39
	3	0,74	0,78	0,83	0,76	0,76	0,78	0,59
	7	0,83	0,74	0,76	0,78	0,76	0,78	0,59

Overall Evaluation Results

After optimizing the parameters using validation kappa values in the training set; the methodology is evaluated using the test set. The methodologies used in the experiment are given in Table 4-14.

Table 4-14 : The signal processing techniques used to classify the CSP features extracted from the BCI Competition IV data

Spatial filter (SF):	Common Average Reference (CAR) filter no (NO) filter
Temporal filter (TF):	8-30 Hz bandpass filter (YES) no (NO) filtering
Feature Normalization (FN):	Linear Feature Normalization (LFN) Gaussian Feature Normalization (GFN) No (NO) Normalization
Feature Vector Normalization (FVN):	Unit-norm Feature Vector Normalization (UFVN)
Classification:	SVM

For evaluation Cohen's kappa coefficient is used as desired in the competition. In Table 4-15, the kappa coefficient values are given for each subject under different filtering, and normalization conditions.

The temporal filtering operation increases the kappa coefficient as it is seen in Table 4-15. The spatial filtering method CAR also increases the performance slightly. If the normalization techniques are considered, only one linear normalization seems to be sufficient. This may be either UFN or LFN.

Under UFN and CAR conditions, the kappa coefficient value for subject A05 is 0.15 while it is 0.59 for the subject A09. This shows the subject dependency of the BCI procedures.

Table 4-15 : The kappa coefficient values for different subject, filtering, and normalization types.

SF	TF	FVN	FN	A01	A02	A03	A04	A05	A06	A07	A08	A09	AVG
CAR	NO	NO	GFN	0,39	0,14	0,48	0,16	0,00	0,12	0,02	0,46	0,35	0,24
CAR	NO	NO	LFN	0,43	0,20	0,50	0,20	0,07	0,14	0,28	0,50	0,45	0,31
CAR	NO	NO	NO	0,43	0,20	0,51	0,18	0,01	0,15	0,29	0,51	0,52	0,31
CAR	NO	UFVN	GFN	0,27	0,00	0,09	0,00	0,00	0,03	0,00	0,28	0,05	0,08
CAR	NO	UFVN	LFN	0,42	0,18	0,51	0,16	0,01	0,18	0,30	0,53	0,52	0,31
CAR	NO	UFVN	NO	0,43	0,20	0,53	0,20	0,08	0,15	0,28	0,53	0,51	0,32
CAR	YES	NO	GFN	0,45	0,17	0,48	0,34	0,07	0,17	0,41	0,48	0,50	0,34
CAR	YES	NO	LFN	0,56	0,25	0,57	0,41	0,18	0,22	0,44	0,54	0,58	0,42
CAR	YES	NO	NO	0,56	0,26	0,55	0,40	0,15	0,22	0,44	0,54	0,59	0,41
CAR	YES	UFVN	GFN	0,01	0,01	0,01	0,00	0,00	0,02	0,11	0,14	0,26	0,06
CAR	YES	UFVN	LFN	0,56	0,24	0,57	0,35	0,14	0,21	0,52	0,52	0,57	0,41
CAR	YES	UFVN	NO	0,56	0,26	0,56	0,39	0,15	0,22	0,46	0,55	0,59	0,42
NO	NO	NO	GFN	0,39	0,10	0,47	0,19	0,02	0,12	0,28	0,44	0,50	0,28
NO	NO	NO	LFN	0,41	0,15	0,51	0,21	0,07	0,15	0,27	0,49	0,47	0,30
NO	NO	NO	NO	0,42	0,14	0,52	0,21	0,06	0,16	0,28	0,52	0,51	0,31
NO	NO	UFVN	GFN	0,33	0,07	0,36	0,13	0,00	0,08	0,23	0,35	0,39	0,21
NO	NO	UFVN	LFN	0,40	0,15	0,51	0,20	0,01	0,17	0,28	0,51	0,51	0,31
NO	NO	UFVN	NO	0,39	0,18	0,52	0,21	0,06	0,16	0,26	0,50	0,53	0,31
NO	YES	NO	GFN	0,44	0,14	0,47	0,30	0,10	0,23	0,46	0,54	0,47	0,35
NO	YES	NO	LFN	0,55	0,22	0,55	0,36	0,14	0,23	0,45	0,54	0,58	0,40
NO	YES	NO	NO	0,53	0,24	0,54	0,39	0,13	0,24	0,51	0,55	0,59	0,41
NO	YES	UFVN	GFN	0,41	0,08	0,37	0,20	0,04	0,18	0,32	0,53	0,48	0,29
NO	YES	UFVN	LFN	0,53	0,15	0,56	0,32	0,13	0,22	0,49	0,54	0,53	0,38
NO	YES	UFVN	NO	0,54	0,22	0,56	0,37	0,12	0,23	0,54	0,54	0,61	0,41

4.2.3 Conclusion

The maximum kappa value obtained in this experiment is given with results of the other research groups attended the competition in Table 4-16. The average of the kappa values obtained for each subject is comparable with the other results. The maximum kappa value is obtained for subject A09. This kappa value is almost equal to the best two results in the competition.

Table 4-16 : The results of BCI competition IV

group	Avg.	Subject								
		A01	A02	A03	A04	A05	A06	A07	A08	A09
1	0.57	0.68	0.42	0.75	0.48	0.40	0.27	0.77	0.75	0.61
2	0.52	0.69	0.34	0.71	0.44	0.16	0.21	0.66	0.73	0.69
METU BCI	0,42	0,56	0,25	0,57	0,41	0,18	0,22	0,44	0,54	0,58
3	0.31	0.38	0.18	0.48	0.33	0.07	0.14	0.29	0.49	0.44
4	0.30	0.46	0.25	0.65	0.31	0.12	0.07	0.00	0.46	0.42
5	0.29	0.41	0.17	0.39	0.25	0.06	0.16	0.34	0.45	0.37

4.3 METU Brain Research Laboratory BCI Experiments

Depending on the experiments performed on BCI competition III dataset V and BCI competition IV dataset IIa, CSP based SVM classification found suitable to be used in BCI studies due to accuracy and performance considerations. Therefore, it has been the main classification algorithm used in Brain Research Laboratory BCI experiments. These experiments can be grouped into to as offline and online experiments. In the offline experiments it is aimed to find suitable number and type of tasks to control a motor imagery based BCI. Then, the application developed depending on the offline experiment results is tested in the online experiment.

4.3.1 Offline Experiments

In the offline experiments conducted in METU Brain Research Laboratories, it is aimed to determine the number and types of motor imagery tasks for a specific subject (subject A) to control a BCI. For that reason five different motor imagery tasks are studied. These tasks are;

1. Imagination of tongue movements,
2. Imagination of left hand movements,
3. Imagination of right hand movements,
4. Imagination of left foot movements,
5. Imagination of right foot movements.

4.3.1.1 Explanation of the Experiment

The subject A, a 27 year old male, attended 5 runs separated by short breaks on the same day. One run consists of 60 trials (12 for each of the five possible classes), yielding a total of 300 trials.

In the experiments, the subject A was sitting relaxed in an ordinary armchair in front of a computer screen. The timing scheme of the trials is the same with the BCI Competition IV dataset 2a (see section 4.2.1). The same paradigm is extended to five classes. The images representing each class and the fixation cross used in the trials are given in Figure 4-7.

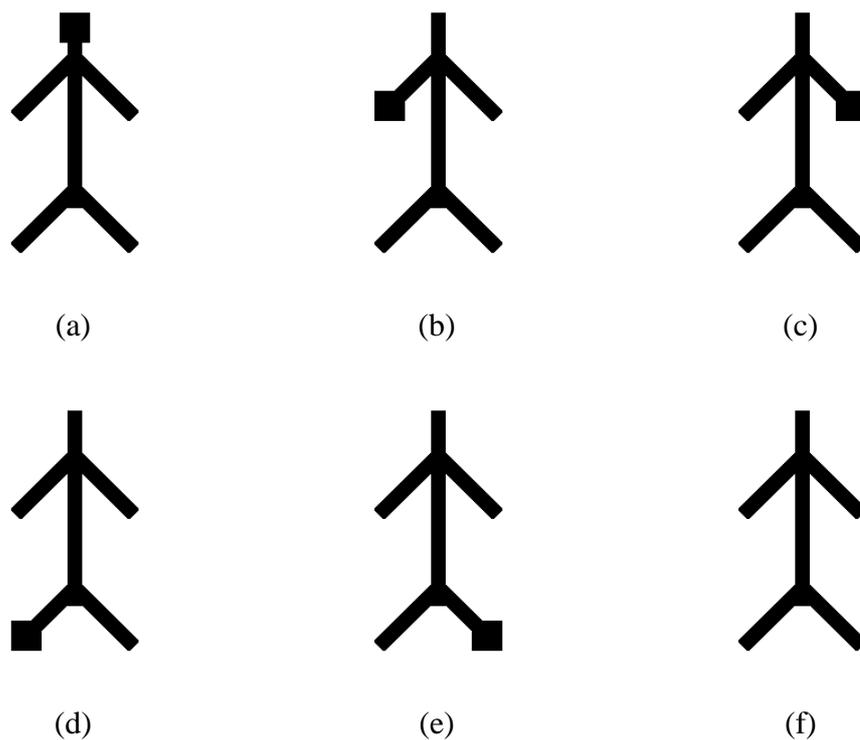


Figure 4-7 : The images representing (a) tongue, (b) left hand, (c) right hand, (d) left foot, (e) right foot movement imagination and (f) the fixation cross used in the trials.

The data acquisition is performed by 10-channel EEG developed for BCI studies in METU Brain Research Laboratory. The electrodes are placed on the skull with a standard EEG cap by applying conductive gel in order to decrease the contact impedance. The montage of the electrodes is given in Figure 4-8 according to the 10-20 electrode system.

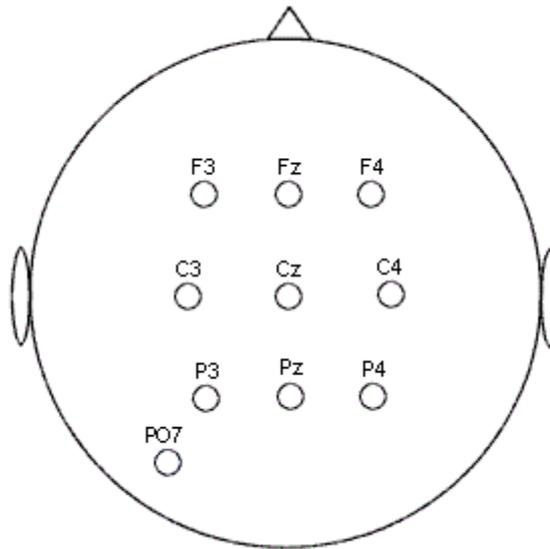


Figure 4-8 : Electrode configuration used in the experiment.

All signals were recorded monopolarly with the mastoids serving as reference. The signals were band-pass filtered between 0.1Hz and 40 Hz via the analog hardware. Also, an additional 50Hz analog notch filter was enabled to suppress the line noise. Then, the signals were sampled at 1000Hz with an amplification factor of 10000 [71].

4.3.1.2 Evaluation of the Data

5-fold cross validation is utilized to evaluate the performance in classifying different motor imagery tasks. For that purpose, each run has served as the test data for evaluating the classifier constructed with the remaining runs. The overall performance is calculated taking the average of 5 performance results. The

classification accuracy, Cohen's kappa coefficient, and Nykopp's information transfer are used as performance measures in the experiments are. However; Nykopp's information transfer have been more significant while selecting the suitable number and types of task for an motor imagery based BCI application.

4.3.1.3 Results

In order to optimize the parameters γ , C , and p in CSP based SVM classification, randomly selected 75% of the training data is used to construct a classifier and the remaining 25% is used to validate the performance of the model for the given the $\gamma - C - p$ combination. While dividing the training data into partitions, uniform distribution of the features for different classes is considered. The final classification model is constructed with the parameters giving maximum validation kappa value using whole training data. Then, the performance of the model is evaluated on the evaluation data. The process is repeated 5 times for 5 different evaluation data selections (i.e. for each of 5 runs). The average of the repetitions will be referred as 5-fold cross validation performance. In Table 4-17, the validation accuracy results; in Table 4-18 validation kappa coefficient results; in Table 4-19, validation information transfer results obtained for each validation process with different runs are given. In Table 4-20, Table 4-21, and Table 4-22 the average (5-fold cross validation) accuracy, kappa and information transfer values for different motor imagery task combinations are given respectively. In the motor imagery task combinations, each task is represented with a number. These numbers and corresponding motor imagery tasks are given below.

- 1: Imagination of tongue movements,
- 2: Imagination of left hand movements,
- 3: Imagination of right hand movements,
- 4: Imagination of left foot movements,
- 5: Imagination of right foot movements.

Validation Results:

Table 4-17 : Validation accuracy results obtained for each validation process with different runs.

Run:	1	2	3	4	5	Average
Tasks						
1 2	60,38	69,20	69,77	66,34	74,59	68,06
1 3	63,15	71,00	72,63	78,51	76,14	72,29
1 4	63,64	66,26	77,94	77,37	74,43	71,93
1 5	64,13	75,49	75,11	67,73	78,35	72,16
2 3	51,88	62,50	72,04	65,20	67,97	63,92
2 4	53,84	68,06	46,81	45,34	54,41	53,69
2 5	50,00	55,07	59,25	51,47	67,65	56,69
3 4	51,39	59,48	69,22	62,91	60,87	60,77
3 5	54,00	59,64	68,45	63,48	58,09	60,73
4 5	60,62	64,54	53,54	54,25	64,62	59,51
1 2 3	43,85	52,18	62,13	59,91	60,73	55,76
1 2 4	38,13	58,12	46,02	46,30	56,26	48,97
1 2 5	37,31	46,95	53,05	45,42	62,47	49,04
1 3 4	48,09	52,56	60,28	61,87	56,37	55,84
1 3 5	41,34	51,53	56,00	51,69	55,56	51,22
1 4 5	36,76	53,16	48,35	50,65	51,36	48,06
2 3 4	32,79	49,18	45,83	43,08	42,54	42,68
2 3 5	33,61	38,02	45,21	48,64	44,28	41,95
2 4 5	31,75	43,90	34,79	43,68	46,95	40,22
3 4 5	41,67	49,56	39,16	41,72	40,09	42,44
1 2 3 4	27,29	43,50	42,30	39,13	47,79	40,00
1 2 3 5	32,84	41,59	43,27	44,69	42,57	40,99
1 2 4 5	26,80	42,08	38,21	38,11	45,71	38,18
1 3 4 5	31,74	45,71	39,73	44,04	40,69	40,38
2 3 4 5	27,49	36,52	29,16	32,31	35,83	32,26
1 2 3 4 5	21,73	35,49	34,38	37,75	38,63	33,60
Average	43,32	53,51	53,18	52,37	55,57	51,59

Table 4-18 : Validation kappa coefficient results obtained for each validation process with different runs

Run:	1	2	3	4	5	Average
Tasks						
1 2	0,21	0,38	0,40	0,33	0,49	0,36
1 3	0,26	0,42	0,45	0,57	0,52	0,44
1 4	0,27	0,33	0,56	0,55	0,49	0,44
1 5	0,28	0,51	0,51	0,35	0,57	0,44
2 3	0,04	0,25	0,44	0,30	0,36	0,28
2 4	0,08	0,36	-0,06	-0,09	0,09	0,07
2 5	0,00	0,10	0,16	0,03	0,35	0,13
3 4	0,03	0,19	0,38	0,26	0,22	0,22
3 5	0,08	0,19	0,37	0,27	0,16	0,21
4 5	0,21	0,29	0,06	0,08	0,29	0,19
1 2 3	0,16	0,28	0,43	0,40	0,41	0,34
1 2 4	0,07	0,37	0,19	0,19	0,34	0,23
1 2 5	0,06	0,20	0,29	0,18	0,44	0,23
1 3 4	0,22	0,29	0,40	0,43	0,35	0,34
1 3 5	0,12	0,27	0,34	0,28	0,33	0,27
1 4 5	0,05	0,30	0,23	0,26	0,27	0,22
2 3 4	-0,01	0,24	0,19	0,15	0,14	0,14
2 3 5	0,00	0,07	0,18	0,23	0,16	0,13
2 4 5	-0,02	0,16	0,02	0,16	0,20	0,10
3 4 5	0,13	0,24	0,08	0,13	0,10	0,13
1 2 3 4	0,03	0,25	0,23	0,19	0,30	0,20
1 2 3 5	0,10	0,22	0,24	0,26	0,23	0,21
1 2 4 5	0,02	0,23	0,18	0,17	0,28	0,18
1 3 4 5	0,09	0,28	0,20	0,25	0,21	0,21
2 3 4 5	0,03	0,15	0,05	0,10	0,14	0,10
1 2 3 4 5	0,02	0,19	0,18	0,22	0,23	0,17
Average	0,10	0,26	0,26	0,24	0,30	0,23

Table 4-19: Validation information transfer results obtained for each validation process with different runs

Run:	1	2	3	4	5	Average
Tasks						
1 2	0,03	0,11	0,46	0,09	0,20	0,18
1 3	0,06	0,13	0,45	0,26	0,21	0,22
1 4	0,06	0,09	0,29	0,27	0,22	0,18
1 5	0,16	0,20	0,20	0,09	0,25	0,18
2 3	0,01	0,05	0,16	0,07	0,10	0,08
2 4	0,02	0,11	0,05	0,07	0,01	0,05
2 5	0,02	0,02	0,03	0,00	0,09	0,03
3 4	0,02	0,03	0,11	0,05	0,03	0,05
3 5	0,01	0,03	0,10	0,07	0,02	0,05
4 5	0,03	0,07	0,01	0,01	0,06	0,04
1 2 3	0,06	0,15	0,26	0,26	0,26	0,20
1 2 4	0,03	0,21	0,77	0,22	0,33	0,31
1 2 5	0,10	0,12	0,21	0,14	0,28	0,17
1 3 4	0,09	0,18	0,26	0,31	0,19	0,21
1 3 5	0,06	0,45	0,19	0,32	0,28	0,26
1 4 5	0,04	0,16	0,20	0,24	0,50	0,23
2 3 4	0,09	0,14	0,16	0,10	0,11	0,12
2 3 5	0,08	0,02	0,09	0,14	0,08	0,08
2 4 5	0,07	0,11	0,03	0,05	0,12	0,08
3 4 5	0,07	0,11	0,04	0,05	0,03	0,06
1 2 3 4	0,07	0,29	0,22	0,39	0,25	0,24
1 2 3 5	0,13	0,21	0,22	0,25	0,23	0,21
1 2 4 5	0,07	0,16	0,18	0,36	0,65	0,28
1 3 4 5	0,11	0,24	0,17	0,53	0,40	0,29
2 3 4 5	0,08	0,08	0,14	0,09	0,12	0,10
1 2 3 4 5	0,12	0,35	0,23	0,52	0,26	0,30
	0,06	0,15	0,20	0,19	0,20	0,16

Overall Evaluation Results:

Table 4-20 : Average (5-fold cross validation) classification accuracies for different motor imagery task combinations

	CAR LFN	CAR UFVN	NO LFN	NO UFVN
1 2	63,42	64,54	68,92	68,06
1 3	68,11	66,33	70,75	72,29
1 4	69,33	68,32	71,03	71,93
1 5	69,77	70,29	71,47	72,16
2 3	61,76	62,73	62,35	63,92
2 4	58,12	59,38	54,00	53,69
2 5	59,48	58,66	57,00	56,69
3 4	62,48	62,60	62,99	60,77
3 5	59,18	61,17	59,35	60,73
4 5	57,21	56,66	59,09	59,51
1 2 3	52,87	52,55	54,23	55,76
1 2 4	46,68	46,46	48,80	48,97
1 2 5	48,99	50,25	50,26	49,04
1 3 4	50,32	52,00	54,86	55,84
1 3 5	51,41	51,48	52,79	51,22
1 4 5	47,17	49,96	49,82	48,06
2 3 4	43,17	44,70	40,83	42,68
2 3 5	43,46	44,09	43,49	41,95
2 4 5	39,21	39,96	38,47	40,22
3 4 5	43,17	42,48	41,28	42,44
1 2 3 4	40,03	40,36	40,10	40,00
1 2 3 5	40,70	41,83	39,01	40,99
1 2 4 5	36,48	37,16	37,26	38,18
1 3 4 5	39,69	40,60	40,10	40,38
2 3 4 5	33,13	33,06	30,87	32,26
1 2 3 4 5	32,41	33,06	31,30	33,60
Average:	50,68	51,18	51,17	51,59

Table 4-21 : Average (5-fold cross validation) kappa values for different motor imagery task combinations

	CAR LFN	CAR UFVN	NO LFN	NO UFVN
1 2	0,27	0,29	0,38	0,36
1 3	0,36	0,32	0,42	0,44
1 4	0,39	0,37	0,42	0,44
1 5	0,39	0,41	0,43	0,44
2 3	0,24	0,26	0,25	0,28
2 4	0,16	0,19	0,08	0,07
2 5	0,19	0,17	0,14	0,13
3 4	0,25	0,25	0,26	0,22
3 5	0,18	0,22	0,19	0,21
4 5	0,14	0,13	0,18	0,19
1 2 3	0,29	0,29	0,31	0,34
1 2 4	0,20	0,20	0,23	0,23
1 2 5	0,23	0,25	0,25	0,23
1 3 4	0,25	0,28	0,32	0,34
1 3 5	0,27	0,27	0,29	0,27
1 4 5	0,21	0,25	0,25	0,22
2 3 4	0,15	0,17	0,11	0,14
2 3 5	0,15	0,16	0,15	0,13
2 4 5	0,09	0,10	0,08	0,10
3 4 5	0,15	0,14	0,12	0,13
1 2 3 4	0,20	0,20	0,20	0,20
1 2 3 5	0,21	0,22	0,19	0,21
1 2 4 5	0,15	0,16	0,16	0,18
1 3 4 5	0,20	0,21	0,20	0,21
2 3 4 5	0,11	0,11	0,08	0,10
1 2 3 4 5	0,16	0,16	0,14	0,17
Average:	0,21	0,22	0,22	0,23

Table 4-22 : Average (5-fold cross validation) information transfer values for different motor imagery task combinations

	CAR LFN	CAR UFVN	NO LFN	NO UFVN
1 2	0,07	0,10	0,11	0,18
1 3	0,10	0,12	0,14	0,22
1 4	0,12	0,12	0,16	0,18
1 5	0,14	0,15	0,15	0,18
2 3	0,10	0,07	0,07	0,08
2 4	0,05	0,05	0,07	0,05
2 5	0,05	0,04	0,05	0,03
3 4	0,05	0,05	0,05	0,05
3 5	0,03	0,05	0,04	0,05
4 5	0,02	0,03	0,03	0,04
1 2 3	0,19	0,15	0,18	0,20
1 2 4	0,11	0,15	0,17	0,31
1 2 5	0,25	0,26	0,17	0,17
1 3 4	0,17	0,18	0,20	0,21
1 3 5	0,20	0,20	0,23	0,26
1 4 5	0,14	0,16	0,16	0,23
2 3 4	0,07	0,08	0,10	0,12
2 3 5	0,09	0,15	0,12	0,08
2 4 5	0,08	0,09	0,08	0,08
3 4 5	0,06	0,07	0,06	0,06
1 2 3 4	0,17	0,16	0,19	0,24
1 2 3 5	0,20	0,22	0,26	0,21
1 2 4 5	0,17	0,17	0,19	0,28
1 3 4 5	0,18	0,24	0,24	0,29
2 3 4 5	0,10	0,12	0,09	0,10
1 2 3 4 5	0,19	0,31	0,31	0,30
Average:	0,12	0,13	0,14	0,16

If the average of the accuracy, kappa and information transfer values for different motor imagery task combinations are analyzed in Table 4-20, Table 4-21, Table 4-22; it is seen that the maximum average values are obtained for no spatial filtering and unit-norm feature vector normalization. CAR filtering did not contribute the results as in the experiments on BCI III and IV competition dataset. The improvement that CAR filtering provided the results was biggest in the BCI Competition III experiments. It also provided a slight improvement in the BCI Competition IV experiments. However the results are better without using a CAR filter in this experiment. This may be due to the number of electrodes used in the calculation of the CAR filter. It was 32 in BCI Competition III experiments, 22 in BCI Competition IV experiments and 8 in this experiment. The common average calculated using more number of electrodes gives a more reliable reference.

It is observed in Table 4-20 and Table 4-21 that the average kappa and accuracy values decrease while the number of tasks are increasing as expected. However the case may be different for information transfer. With no spatial filtering and UFVN, which is the best considering the average evaluation criteria values, the information transfer is maximum for the tasks 1, 2, and 4 which are tongue, left hand, right hand movement imaginations respectively. Therefore these tasks have been studied in the online experiments.

4.3.2 Online Experiments

In these experiments a BCI that assist the paralyzed people for controlling environmental devices is designed and a test application for the design is realized. The aim of the design is to basically make it possible for a subject to select items from a menu by the help of motor imagery tasks.

There are several BCI based icon selection applications in the literature. An overview of these applications is given in section 2.4. In this study, a menu of icons is designed in the form of a tree with 2 levels. In the first level, device selection is performed and in the second level the command related to the device is selected. The elements in the first level and the related commands for these elements are summarized in Table 4-23. The first and second levels of the menu are shown in Figure 4-9 and Figure 4-10 respectively.

Table 4-23 : The elements and commands for the menu designed.

Level 1:	Bed	TV	WC	Clima	Light
Level 2:	-Up	-Channel Up		-Temperature Up	-Light On
	-Down	-Channel Down		-Temperature Down	-Light Off
		-Volume up			
		-Volume down			

For each element in the menu, there exists icons representing different motor imagery tasks at the upper left corner of each element. For each element there are two tasks to be performed consecutively. The reason of using multiple tasks is to increase the number of possible selections. The timing of the tasks is synchronized by the icon at the center of upper half of the windows. The motor imagery tasks are performed when this icon is shown. The type and number of tasks are selected according to the offline experiments performed with subject A.

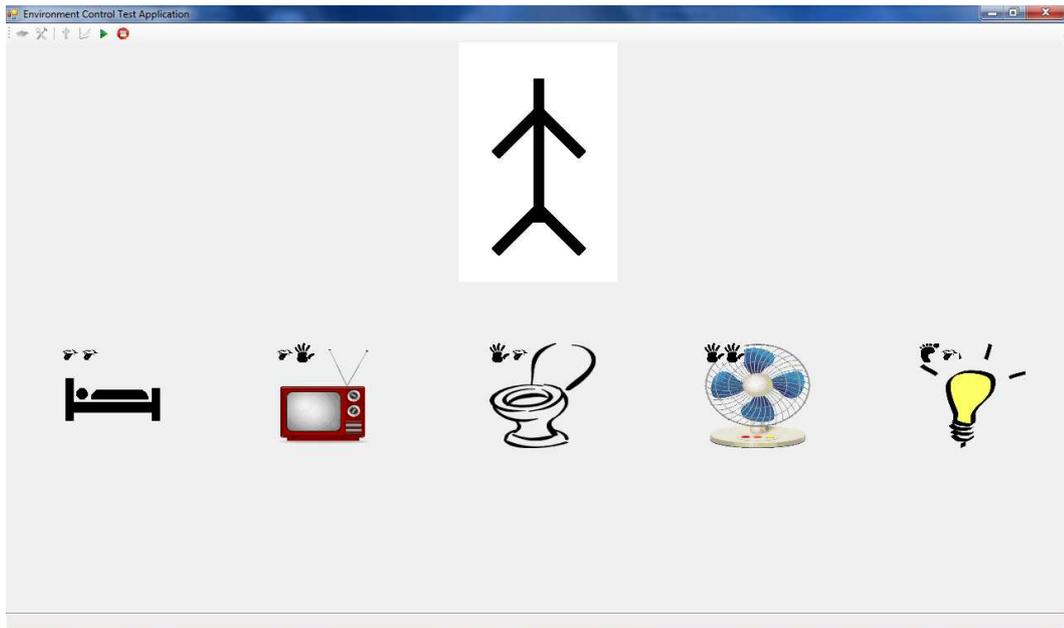


Figure 4-9 : The first level of the menu. The elements are selected by performing the tasks symbolized with small icons at the top left corner of each element.

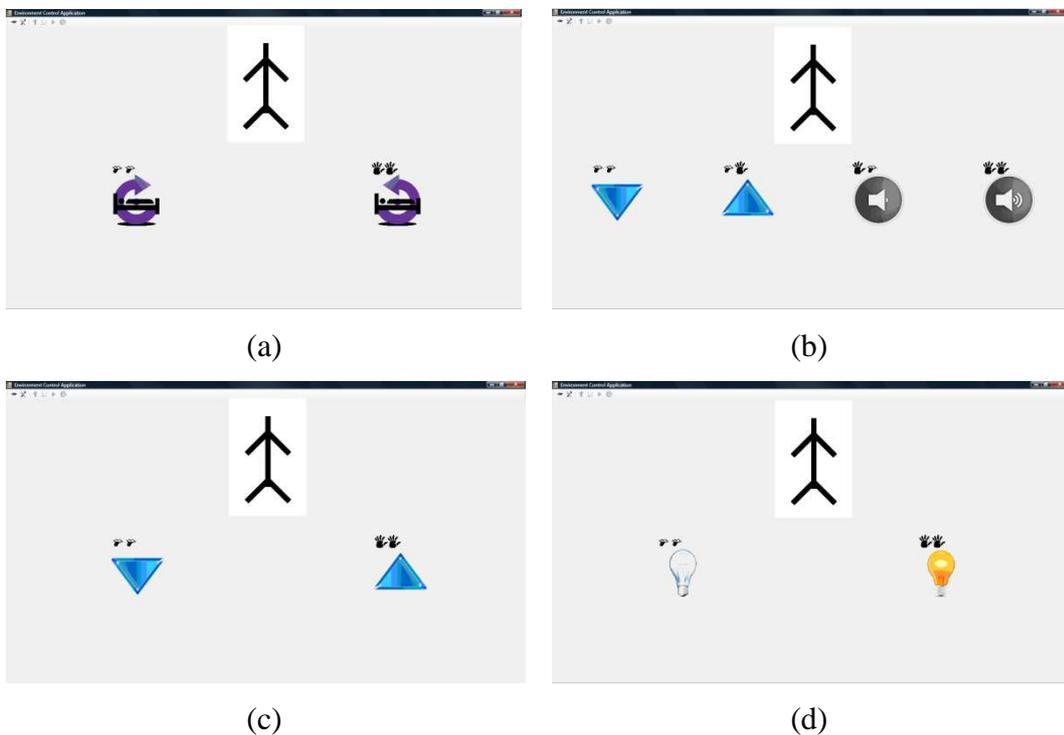


Figure 4-10 The second level of the menu. The commands for (a) motorized bed, (b) TV, (c) air conditioner, (d) light bulb are selected by performing the tasks symbolized with small icons at the top left corner of each element.

To test the interface designed, an application is implemented in Visual Studio .NET 2008 using C# language. The application is integrated with the 10-channel EEG data acquisition system developed for BCI studies in METU Brain Research Laboratory [71] for online experiments. In the experiments, the first level of the design given above is tested. For that purpose, randomly selected one of the five elements in the first level of the menu is presented to the subject at the beginning of each trial. The subject perceives the icons representing the tasks to select the element. Then the element disappears and synchronization icon appears after a short time. The subject performs the imagery task during the synchronization icon is visible. After the imagination, the synchronization icon disappears and the same element appears in order to remind the second task to be performed to select the element. Then the synchronization icon is presented again for the imagination of second task. The time scheme for a single trial is illustrated in Figure 4-12.

In the application, three motor imagery tasks have been used to make selection among 5 icons shown in Figure 4-11. These tasks are selected to be imagination of tongue movements, imagination of left hand movements, and imagination of left foot movements since that combination has provided the best transfer rate in the offline experiments with subject A. The classification model constructed for these tasks in the offline experiments have been used in the online experiments.

The experiment is comprised of 4 runs separated by short breaks. In a single run, each element in Figure 4-11 is presented 4 times in random order. This yields a total of 20 trials per run. The success of the subject to perform the task couple to select the element in the trial is shown in the lower left corner of the interface for each task separately.

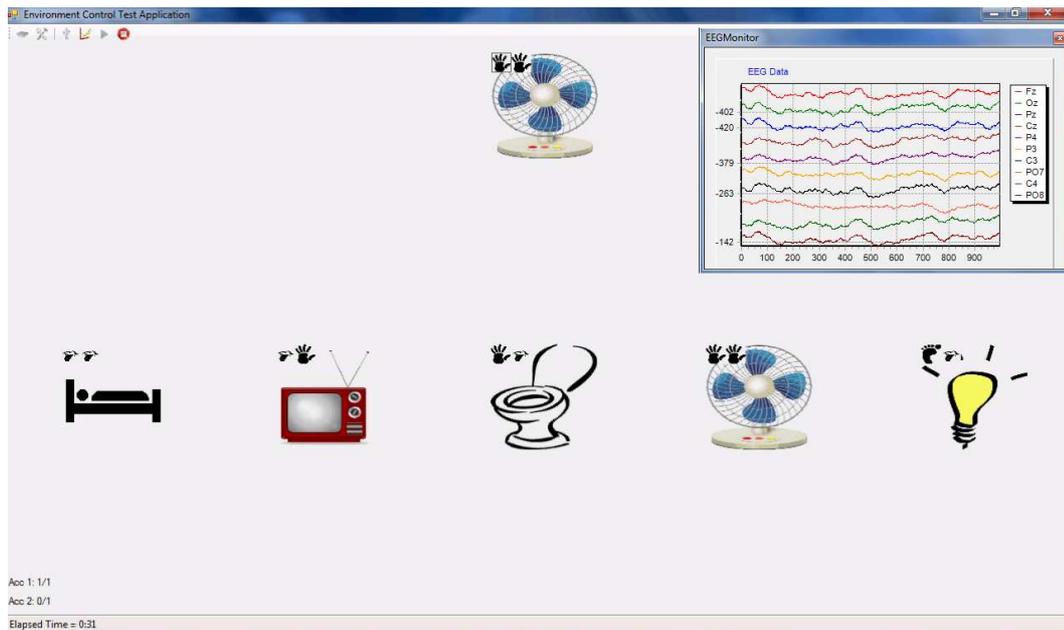


Figure 4-11 : The interface of the test application implemented. The EEG dataflow is checked before and after each run from the window in upper right corner.

The imagination period in each trial can be repeated to increase the classification accuracy. In this experiment 1 and 2 repetitions are tested. More repetitions are not tried since increasing time results in fatigue and concentration loss in the subject. The time scheme of a single trial for N repetitions is given in Figure 4-12.

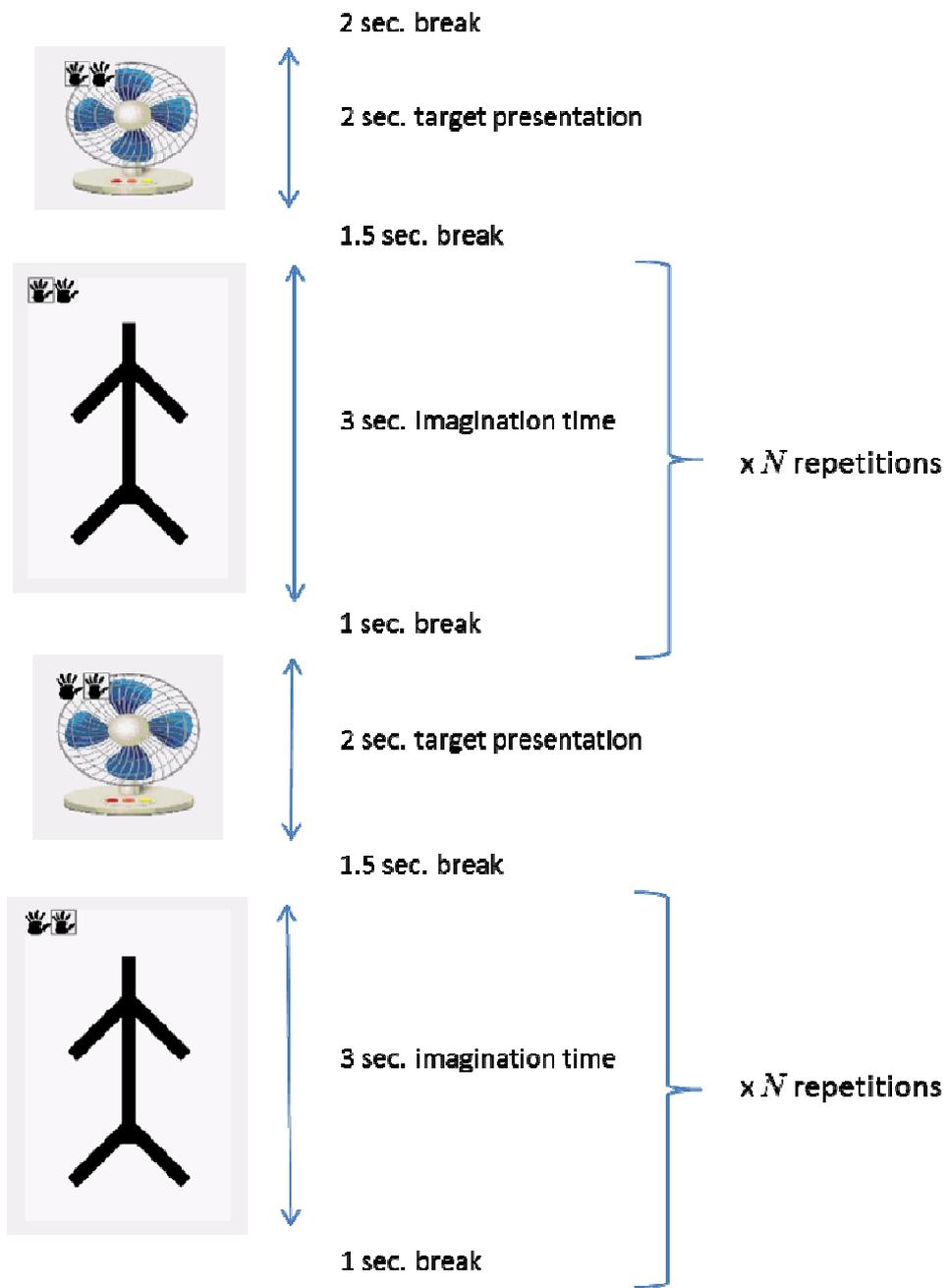


Figure 4-12 : The time scheme of a single trial for N repetitions.

The results for the tasks related to the elements of the menu are recorded separately and given in Table 4-24.

Table 4-24 : The accuracy results of the online experiments.

Run no	Number of repetitions	Accuracy	
		Task 1	Task 2
1	1	6/20	8/20
2	1	9/20	6/20
3	2	13/20	9/20
4	2	7/20	11/20

The random accuracy for selecting one of three tasks is 33.3%. In this experiment, this selection is performed with maximum 65% accuracy in the 2-repetitions case. Although, this accuracy is low in terms of controlling a BCI perfectly, the value may increase with the training of the subject regularly.

CHAPTER 5

CONCLUSION

In this study, different motor imagery tasks are classified in EEG signal using several signal processing techniques. These techniques can be summarized as follows:

- Signal enhancement: band-pass filtering, Small Laplacian (SL) filtering, Large Laplacian (SL) filtering and common average reference (CAR) filtering
- Feature extraction: Common Spatial Pattern (CSP), Power Spectral Density (PSD), Principal Component Analysis (PCA)
- Normalization: Linear Feature Normalization (LFN), Gaussian Feature Normalization (GFN), Unit-norm Feature Vector Normalization (UFVN)
- Classification: Support Vector Machines (SVM), Artificial Neural Networks (ANN)
- Evaluation: classification accuracy, Cohen's kappa coefficient, Nykopp's information transfer

All these methodologies are experimented on the dataset provided for BCI Competition III. In the experiments it is observed that CSP features give better results than PSD for both SVM and ANN classifications. Therefore it is preferred

in the experiments on BCI Competition IV dataset and METU Brain Research Laboratory experiments.

It is also observed that CAR filtering enhances the quality of the EEG signal compared to the no filtering, LL filtering, and SL filtering cases. However, the contribution of the CAR filtering to the classification accuracy decreased with the decreasing number of electrodes. The improvement that CAR filtering provided to the results was highest in the BCI Competition III experiments in which CAR is calculated using 32 electrodes. It also provided a slight improvement in the BCI Competition IV dataset experiments in which 22 electrodes are used. However, the results were better without using a CAR filter in the experiments conducted in METU Brain Research Laboratories with 10 electrodes. The common average calculated using more number of electrodes gave a more reliable reference.

Another observation in the experiments is the effect of temporal filtering. Filtering the EEG signal in the pass band 8-30Hz increased the classification accuracy in the experiments. This is expected since it is the frequency band related to the motor imagery.

The best result is obtained with CSP-based ANN in BCI competition III. An average accuracy of 70,96% is obtained among three subjects. This result is better than the winner of the competition. However the SVM results for CSP features were also comparable with the other results in the competition. Furthermore, the training and testing duration of SVM was observed to be 4-5 times shorter than ANN. This makes SVM preferable in the experiments and online applications considering short response time. The CSP-based SVM is also tested on the BCI Competition IV data which includes EEG of 4 different motor imagery tasks. According to the obtained results, CSP-based SVM has the 3rd rank among the participants of the competition with the average kappa value of 0.42 for all subjects.

After experimenting the methodologies on the data recorded from different subjects for BCI Competition III and IV, some experiments have also been conducted in the METU Brain research laboratory. In the offline part of these experiments, 5 different motor imagery tasks are studied with subject A. Depending on the information transfer calculated for different combinations of these tasks, tongue movement imagination, left hand movement imagination, left foot movement imagination tasks found suitable for subject A to control a motor imagery based BCI. Considering this specific offline experiment, an interface is designed for subject A to control assistive environmental devices. Then, a test application is implemented and online performance of the design is evaluated. The subject A achieved to select one of the three tasks with a maximum accuracy of 65% in one of the runs. Although, this accuracy is low in terms of controlling a BCI perfectly, the value may increase with the training of the subject regularly.

To summarize, the building blocks of a BCI system are studied step by step, focusing on a SMR-based environmental control system in the scope of the thesis. One of the major contributions of the thesis to the literature is the use of multiclass extension of CSP features together with a 3 layer feed-forward ANN to classify SMR. With employing CAR, and 8-30 Hz band-pass filter together with this classification approach, it is achieved to classify 3-class EEG data provided in BCI Competition III with an average accuracy of 70,96%. This result is 2,31% better than the winner of the competition. Another contribution is the SMR-based BCI design for controlling assistive environmental devices. In the early online experiments to evaluate the potential of the design, promising results are obtained.

REFERENCES

- [1] R. Sitaram, A. Caria, R. Veit, T. Gaber, G. Rota, A. Kuebler, and N. Birbaumer, "fMRI brain-computer interface: A tool for neuroscientific research and treatment", *Computational intelligence and neuroscience*, Jan. 2007.
- [2] S.M. Coyle, T.E. Ward, and C.M. Markham, "Brain-computer interface using a simplified functional near-infrared spectroscopy system", *Journal of neural engineering*, vol. 4, Sep. 2007, pp. 219-226.
- [3] N.J. Hill, T.N. Lal, M. Schr, T. Hinterberger, G. Widman, C.E. Elger, B. Sch, and N. Birbaumer, "Classifying Event-Related Desynchronization in EEG , ECoG and MEG signals", *Interface*, 2006, pp. 404-413
- [4] J. a Pineda, "The functional significance of mu rhythms: translating "seeing" and "hearing" into "doing"", *Brain research reviews*, vol. 50, 2005, pp. 57-68.
- [5] J.J. Daly and J.R. Wolpaw, "Brain computer interfaces in neurological rehabilitation", *The Lancet Neurology*, vol. 7, 2008, pp. 1032-1043.
- [6] 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, 2009, pp. 187-199.
- [7] M.V. Gerven, J. Farquhar, R. Schaefer, R. Vlek, J. Geuze, A. Nijholt, N. Ramsey, P. Haselager, L. Vuurpijl, S. Gielen, and P. Desain, "The brain computer interface cycle", *Journal of Neural Engineering*, vol. 6, 2009.
- [8] E. Niedermeyer and F.H.L.D. Silva, "Electroencephalography: basic principles, clinical applications, and related fields", *Lippincott Williams & Wilkins*, 2005.
- [9] "Electroencephalography," <http://en.wikipedia.org/wiki/Eeg>, last visited on 15/01/2011.
- [10] "Intendix," <http://www.intendix.com/>, last visited on 15/01/2011.
- [11] F. Lotte and M. Congedo, "A review of classification algorithms for EEG-based brain computer interfaces", *Journal of neural engineering* , vol. 4, 2007

- [12] N. Birbaumer, "Breaking the silence: Brain-computer interfaces (BCI) for communication and motor control," *Psychophysiology*, vol. 43, pp. 517-532.
- [13] J.S. Brumberg, A. Nieto-Castanon, P.R. Kennedy, and F.H. Guenther, "Brain Computer Interfaces for Speech Communication", *Speech communication*, vol. 52, Apr. 2010, pp. 367-379.
- [14] "medGadget - internet journal of emerging medical technologies" http://medgadget.com/archives/2007/03/the_first_comme_1.html, last visited on 15/01/2011.
- [15] "Precision Design Laboratory," <http://www.mech.utah.edu/~bamberg/>, last visited on 15/01/2011.
- [16] J.A. Wilson, E.A. Felton, P.C. Garell, G. Schalk, and J.C. Williams, "ECoG factors underlying multimodal control of a brain-computer interface", *IEEE Trans. Neural Syst. Rehab. Eng.*, vol. 14, pp. 246-250.
- [17] "Magnetoencephalography," <http://en.wikipedia.org/wiki/Magnetoencephalography>, last visited on 15/01/2011.
- [18] G. Pires and U. Nunes, "A Brain Computer Interface Methodology Based on a Visual P300 Paradigm", *IEEE/RSJ International Conference on Intelligent Robots and Systems Control*, 2009, pp. 4193-4198.
- [19] O. Friman, I. Volosyak, and A. Gräser, "Multiple channel detection of steady-state visual evoked potentials for brain-computer interfaces.," *IEEE transactions on biomedical engineering*, vol. 54, Apr. 2007, pp. 742-750.
- [20] N. Birbaumer, "Slow Cortical Potentials : Plasticity , Operant Control , and Behavioral Effects", *The Neuroscientist*, vol. 5, 2009, pp. 74-78.
- [21] G.G. Molina, T. Tsoneva, A. Nijholt, "Emotional Brain-Computer Interfaces" *3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, 2009, p. 1-9.
- [22] G. Pfurtscheller and C. Neuper, "Motor Imagery and Direct Brain Computer Communication" *Proceedings of the IEEE*, vol. 89, 2001, pp. 1123-1134.
- [23] N. Birbaumer, A. Kübler, N. Ghanayim, T. Hinterberger, J. Perelmouter, J. Kaiser, I. Iversen, B. Kotchoubey, N. Neumann, and H. Flor, "The Thought Translation Device (TTD) for Completely Paralyzed Patients" *IEEE transactions on rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 8, 2000, pp. 190-193.

[24] J.R. Wolpaw and D.J. Mcfarland, "Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans" *Proc. Nat. Acad. Sci USA*, vol. 101, pp. 17849-17854.

[25] L. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials", *Electroencephalography and Clinical Neurophysiology*, vol. 70, Dec. 1988, pp. 510-523.

[26] A. Com, N.V. Manyakov, N. Chumerin, J. a K. Suykens, and M.M.V. Hulle, "Feature Extraction and Classification of EEG Signals for Rapid P300 Mind Spelling," *2009 International Conference on Machine Learning and Applications*, Dec. 2009, pp. 386-391.

[27] "P3SpellerTask," http://www.bci2000.org/wiki/index.php/User_Reference:P3SpellerTask, last visited on 15/01/2011.

[28] M.G.M. Cherubini and F. Babiloni, "Non-invasive brain-computer interface system: Towards its application as assistive technology", *Brain Research Bulletin*, vol. 75, pp. 796-803.

[29] J.K. Chapin, K.A. Moxon, R.S. Markowitz, and M.A. Nicolelis, "Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex", *Nature neuroscience*, vol. 2, Jul. 1999, pp. 664-670.

[30] M. Velliste, S. Perel, M.C. Spalding, A.S. Whitford, and A.B. Schwartz, "Cortical control of a prosthetic arm for self-feeding", *Nature*, vol. 453, Jun. 2008, pp. 1098-1101.

[31] G. Pfurtscheller, C. Guger, G. Müller, G. Krausz, and C. Neuper, "Brain oscillations control hand orthosis in a tetraplegic", *Neuroscience letters*, vol. 292, Oct. 2000, pp. 211-214.

[32] G.R. Müller-Putz, R. Scherer, G. Pfurtscheller, and R. Rupp, "EEG-based neuroprosthesis control: a step towards clinical practice", *Neuroscience letters*, vol. 382, 2005, pp. 169-174.

[33] R. Ortner, B. Allison, G. Korisek, H. Gaggl, and G. Pfurtscheller, "An SSVEP BCI to Control a Hand Orthosis for Persons With Tetraplegia", *IEEE transactions on neural systems and rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, Sep. 2010.

- [34] M. Palankar, K.J.D. Laurentis, R. Alqasemi, E. Veras, R. Dubev, Y. Arbel, E. Donchin, "Control of a 9-DoF Wheelchair-Mounted Robotic Arm System Using a P300 Brain Computer Interface : Initial Experiments", *Proceedings of the 2008 IEEE International Conference on Robotics and Biomimetics*, 2009, pp. 348-353.
- [35] 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, 1997, pp. 386-394.
- [36] W.D. Penny, S.J. Roberts, E.A. Curran, and M.J. Stokes, "EEG-based communication: a pattern recognition approach", *IEEE transactions on rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 8, Jun. 2000, pp. 214-215.
- [37] G. Pfurtscheller, C. Neuper, A. Schlögl, and K. Lugger, "Separability of EEG signals recorded during right and left motor imagery using adaptive autoregressive parameters", *IEEE transactions on rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 6, Sep. 1998, pp. 316-325.
- [38] T. Wang, J. Deng, and B. He, "Classifying EEG-based motor imagery tasks by means of time-frequency synthesized spatial patterns", *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, vol. 115, Dec. 2004, pp. 2744-2753.
- [39] L. Qin, L. Ding, and B. He, "Motor imagery classification by means of source analysis for brain-computer interface applications", *Journal of neural engineering*, vol. 1, Sep. 2004, pp. 135-141.
- [40] B. Kamousi, Z. Liu, and B. He, "Classification of motor imagery tasks for brain-computer interface applications by means of two equivalent dipoles analysis", *IEEE transactions on neural systems and rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 13, Jun. 2005, pp. 166-171.
- [41] M. Congedo, F. Lotte, and A. Lécuyer, "Classification of movement intention by spatially filtered electromagnetic inverse solutions", *Physics in medicine and biology*, vol. 51, Apr. 2006, pp. 1971-89.
- [42] K.K. Ang, Z.Y. Chin, H. Zhang, and C. Guan, "Robust filter bank common spatial pattern (RFBCSP) in motor-imagery-based brain-computer interface", *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Jan. 2009, pp. 578-81.

- [43] T. Yan, T. Jingtian, and G. Andong, "Multi-Class EEG Classification for Brain Computer Interface based on CSP", *International Conference on BioMedical Engineering and Informatics*, 2008, pp. 469-472.
- [44] E.B. Sadeghian and M.H. Moradi, "Continuous detection of motor imagery in a four-class asynchronous BCI", *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Jan. 2007, pp. 3241-3244.
- [45] C. Lin and M. Hsieh, "Classification of mental task from EEG data using neural networks based on particle swarm optimization", *Neurocomputing*, vol. 72, Jan. 2009, pp. 1121-1130.
- [46] F. Galán, M. Nuttin, E. Lew, P.W. Ferrez, G. Vanacker, J. Philips, and J.D.R. Millán, "A brain-actuated wheelchair: asynchronous and non-invasive Brain-computer interfaces for continuous control of robots", *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, vol. 119, Sep. 2008, pp. 2159-2169.
- [47] G. Pfurtscheller, C. Neuper, C. Guger, W. Harkam, H. Ramoser, a Schlögl, B. Obermaier, and M. Pregenzer, "Current trends in Graz Brain-Computer Interface (BCI) research" *IEEE transactions on rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 8, Jun. 2000, pp. 216-219.
- [48] R. Palaniappan, R. Paramesran, S. Nishida, and N. Saiwaki, "A new brain-computer interface design using fuzzy ARTMAP" *IEEE transactions on neural systems and rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 10, Sep. 2002, pp. 140-8.
- [49] L. Ke and R. Li, "Classification of EEG Signals by Multi-Scale Filtering and PCA", *IEEE International Conference on Intelligent Computing and Intelligent Systems*, 2009, pp. 362-366.
- [50] H. Lee and S. Choi, "PCA-based linear dynamical systems for multichannel EEG classification" *Proceedings of the 9th International Conference on Neural Information Processing*, vol. 2, pp. 745-749.
- [51] G. Dornhege, B. Blankertz, G. Curio, and K.R. Muller, "Increase information transfer rates in BCI by CSP extension to multi-class", *Advances in Neural Information Processing Systems*, vol. 16, 2004, pp. 733-740.
- [52] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement" *IEEE transactions on rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 8, 2000, pp. 441-446.

- [53] L. Guangquan, H. Gan, and Z. Xiangyang, http://www.bbc.de/competition/iv/results/ds2a/LiuGuangquan_desc.txt, last visited on 15/01/2011.
- [54] “Spectral density,” http://en.wikipedia.org/wiki/Spectral_density, last visited on 15/01/2011.
- [55] “Welch’s method,” http://en.wikipedia.org/wiki/Welch_method, last visited on 15/01/2011.
- [56] K. Pearson, “On lines and planes of closest fit to systems of points in space”, *Philosophical Magazine*, 1901, pp. 559-572.
- [57] S.D. Duda R., Hart P., *Pattern Classification*, 2000.
- [58] F. Lotte, M. Congedo, a Lécuyer, F. Lamarche, and B. Arnaldi, “A review of classification algorithms for EEG-based brain-computer interfaces” *Journal of neural engineering*, vol. 4, Jun. 2007, p. R1-R13.
- [59] C.-wei Hsu, C.-chung Chang, and C.-jen Lin, “A Practical Guide to Support Vector Classification,” *Bioinformatics*, vol. 1, 2010, pp. 1-16.
- [60] “LIBSVM - A Library for Support Vector Machines,” <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>, last visited on 15/01/2011.
- [61] “BCI Competition III,” <http://www.bbc.de/competition/iii/>, last visited on 15/01/2011.
- [62] “BCI Competition IV,” <http://www.bbc.de/competition/iv/>, last visited on 15/01/2011.
- [63] A. Schl, J.E. Huggins, and S.G. Mason, “Evaluation Criteria for BCI Research”, MIT Press Online Computer and Information Sciences Library. Available at: <http://www.citeulike.org/user/jmetzen/article/3211329>
- [64] T. Nykopp, “Statistical Modelling Issues for The Adaptive Brain Interface” *MSc. Thesis*, Department of Electrical and Communications Engineering, Helsinki University of Technology, 2001.
- [65] J.R. Millan, “On the need for on-line learning in brain-computer interfaces” IEEE International Joint Conference on Neural Networks, 2004, pp. 2877-2882.
- [66] “Biosemi Products”, <http://www.biosemi.com/products.htm>, last visited on 15/01/2011.

- [67] G.H. Klem, H.O. Lüders, H.H. Jasper, and C. Elger, “The ten-twenty electrode system of the International Federation of Clinical Neurophysiology”, *Electroencephalography and clinical neurophysiology*. Supplement, vol. 52, Jan. 1999, pp. 3-6.
- [68] C.J.C. Burges, “A Tutorial on Support Vector Machines for Pattern Recognition”, *Data Mining and Knowledge Discovery*, vol. 167, 1998, pp. 121-167.
- [69] C. Brunner and R. Leeb, BCI Competition 2008 – Graz data set A, 2008.
- [70] A. Kai Keng, C. Zheng Yang, W. Chuanchu, G. Cuntai, Z. Haihong, P. Kok Soon, H. Brahim, and T. Keng Peng, BCI Competition IV Dataset 2a Submission by Institute for Infocomm Research, Agency for Science, Technology and Research Singapore (I2R, A*STAR), 2008.
- [71] H. B. Erdoğan, “A Design and Implementation of P300 Based Brain Computer Interface”, *MSc. Thesis*, Department of Electrical and Electronics Engineering, METU, 2009.