INSTITUTO SUPERIOR TÉCNICO
Universidade Técnica de Lisboa

Interface Cérebro-Computador (BCI) no Paradigma de Imagiologia Motora

Carlos Alexandre Pereira Carreiras
(Licenciado)

Dissertação para obtenção do grau de Mestre em

Engenharia Biomédica

Júri

Presidente:	Prof.	Paulo Jorge Peixeiro de Freitas
Orientador:	Prof.	João Miguel Raposo Sanches
Co-Orientador:	Prof.	Luís Henrique Martins Borges de Almeida
Vogal:	Prof.	José Manuel Bioucas Dias
Vogal:	Prof. Dr.	Mamede de Carvalho

Outubro 2011

Acknowledgments

I would like to thank my supervisor, Prof. João Sanches, for his helpful guidance throughout the development of this thesis. Also, I would like to acknowledge my co-supervisor, Prof. Luís Borges de Almeida, for pointing me in the right direction related to the use of the Phase-Locking Factor.

This work would not be possible without the collaboration of the Centro de Electroencefalografia e Neurofisiolofia Clínica and in particular Prof. Dr. Teresa Paiva, who provided the equipment and venue for the EEG acquisitions. Also, a big thank you to all the 6 subjects who submitted their brains for this work, and to David Belo, who helped prepare the subjects for the acquisition.

Abstract

Brain-Computer Interface (BCI) systems attempt to create a direct channel of communication between the brain and a computer. This is especially important for patients that are ”locked in”, as they have limited motor function. In a BCI system, measures of brain activity are used to command a computer application. The most common measure of brain activity used is the Electroencephalogram (EEG). The users, in order to operate the BCI, must acquire conscious control over their brain activity, producing visible changes in the EEG. Here, this is done is through the imagination of motor tasks. In particular, motor tasks produce the so called Event-Related Desynchronization (ERD), which is a result of increased activity in the motor cortex. Traditionally,ERD is measured through the estimation of EEG signal power in specific frequency bands. This describes the first type of features explored in this thesis. The second type is inspired by the fact that neuronal populations in the motor cortex become desynchronized when processing a motor task. The proposed method is based on the phase information from the EEG channels, using the Phase-Locking Factor (PLF). Both feature types were tested in real data obtained from 6 voluntary subjects, who performed 7 motor tasks in an EEG session. The features were classified using Support Vector Machine (SVM) classifiers organized in a hierarchical structure. For the PLF features, additional classifiers were used to evaluate the best combination of motor tasks to control a four-channel BCI system (for instance, to operate a 2D cursor). The results show that the PLF features are better, with an average accuracy of $\approx 86\%$, against an accuracy of $\approx 70\%$ for the band power features. There are no statistically significant differences in accuracy for the different combinations of motor tasks, not being possible to infer with certainty which combination is best. Despite this, slightly better results were obtained for a system using the following motor tasks: movement of the right and left hands, of the left foot and of the right leg. Although more research is still needed, the PLF measure shows promising results for use in a BCI system.

Keywords

Brain-Computer Interface, Imagined Motor Tasks, Electroencephalogram, Event-Related Desynchronization, Band Power Features, Phase-Locking Factor, Support Vector Machine.

Resumo

As Interfaces Cérebro-Computador (BCI) estabelecem um canal de comunicação directo entre o cérebro e um computador. Isto tem particular importância para pacientes que apresentem função motora limitada. Num sistema BCI, mediadas da actividade cerebral são usadas para controlar uma aplicação de computador. Mais frequentemente, é usado o Electroencefalograma (EEG). Os utilizadores, para controlarem a BCI, têm de modular conscientemente a sua actividade cerebral, produzindo alterações visíveis no EEG. Aqui, isto é feito recorrendo à imaginação de tarefas motoras, que originam os chamados Eventos de Dessincronização (ERD), resultado do aumento de actividade no córtex motor. Tradicionalmente, a ERD é medida através da estimação da energia do EEG em bandas de frequência específicas. Isto constitui o primeiro tipo de *features* usadas nesta tese. O segundo tipo inspira-se no facto de populações neuronais no córtex motor perderem sincronia durante o processamento de tarefas motoras. O método proposto baseia-se na fase obtida do EEG, usando o Phase-Locking Factor (PLF). Ambos os tipos de *features* foram testados em dados reais obtidos de 6 voluntários, que executaram 7 tarefas motoras numa sessão de EEG. As *features* foram classificadas usando Support Vector Machine (SVM) organizados numa estrutura hierárquica. Para as *features* extraídas usando o PLF, classificadores adicionais foram utilizados para determinar qual a melhor combinação de tarefas para controlar uma BCI de quatro canais. Os resultados mostram que as *features* PLF são melhores, com uma exactidão média de $\approx 86\%$, contra uma exactidão de $\approx 70\%$ para as *features* baseadas na energia. As diferentes combinações de tarefas não mostram diferenças estatisticamente significativas na precisão, não sendo possível concluir qual a melhor. Não obstante, melhores resultados foram obtidos para a seguinte combinação de tarefas: movimento das mãos esquerda e direita, do pé esquerdo e da perna direita. Apesar de ser necessária mais investigação, as *features* baseadas no PLF mostram ser prometedoras para a sua utilização em sistemas BCI.

Palavras Chave

Interface Cérebro-Computador, Imaginação de Tarefas Motoras, Electroencefalograma, Eventos de Dessincronização, *Features* de Energia, Phase-Locking Factor, Support Vector Machine.

Contents

List of Figures

List of Tables

Abbreviations

ALS Amyotrophic Lateral Sclerosis

AR Auto-Regressive

BCI Brain-Computer Interface

BSS Blind Source Separation

CNS Central Nervous System

CAR Common Average Reference

CSP Common Spatial Patterns

CBP Complex Band Power

DFT Discrete Fourier Transform

ECoG Electrocorticogram

EEG Electroencephalogram

EMG Electromyogram

ENIAC Electronic Numerical Integrator and Computer

EOG Electrooculogram

ERD Event-Related Desynchronization

ERS Event-Related Synchronization

FFT Fast Fourier Transform

FDA Fisher's Discriminant Analysis

FT Fourier Transform

fMRI functional Magnetic Resonance Imaging

GA Genetic Algorithm

HMM Hidden Markov Model

ICA Independent Component Analysis

LOOCV Leave-One-Out Cross Validation

LDA Linear Discriminant Analysis

MEG Magnetoencephalogram

NIRS Near-Infrared Spectroscopy

NN Neural Network

PNS Peripheral Nervous System

PCI Phase Clustering Index

PLF Phase-Locking Factor

PMC Primary Motor Cortex

SL Surface Laplacian

SVM Support Vector Machine

1

Introduction

Contents

1.1 Motivation

The interaction between humans and computers has always been an arduous and difficult task. The first general-purpose electronic computer, the Electronic Numerical Integrator and Computer (ENIAC) developed in 1946, used as input a punched card reader. Two years later, in 1948, the card reader concept was combined with the typewriter, giving birth to an electromechanically controlled typewriter, which eventually evolved to the current keyboard. Later on, the introduction of the computer screen and the development of Graphical User Interfaces required a spatial input method, originating the electronic mouse. Through these technical advances, along with the ever increasing computational power and memory and ever decreasing size, the computer has become an ubiquitous device, providing its users a facilitated and improved communication channel. Indeed, the ability to communicate with other persons is one of the main factors making the life of any human being enjoyable [1].

These traditional input methods, although simple and reliable, aim at adapting the interface to our natural senses or to replace them [2], providing a somewhat indirect route to the user's true intent. Additionally, some users lack the abilities to operate the mentioned input methods. Users suffering from severe motor disabilities such as Amyotrophic Lateral Sclerosis (ALS), spinal-cord injury, stroke and other neuromuscular diseases or injuries have many difficulties carrying out normal day-to-day tasks, not to mention operating a computer, given that these patients cannot control their own muscles. So why not go directly to the source of all thought and action: the brain? Can we extract enough information from the brain to create a new channel of communication between humans and machines? Could that channel be used to improve the quality of life of motor disabled patients? These and other questions are the main focus of current research in Brain-Computer Interface (BCI), a method through which measurements of electric, magnetic or other physical manifestations of brain activity are translated into commands for a computer or other devices [1].

1.2 Main Contributions

The work presented here is the result of a MSc thesis developed over the course of six months, from March 2011 to September 2011. The main goal of this thesis is to develop an Electroencephalogram Brain-Computer Interface based on the imagination of motor tasks, such as moving the right or left hand. The fundamental aspect explored in this work is the neurophysiological aspects of motor tasks that can be identified through the Electroencephalogram (EEG). The focus is given to Event-Related Desynchronization (ERD) and Event-Related Synchronization (ERS), which are, traditionally, identified by measuring the power of specif bands of the EEG. An alternative method, inspired by Almeida *et al.* in [3], is evaluated here. This method, from the field of synchronization quantification, encompasses the definition of a synchronization measure between two signals, the Phase-Locking Factor (PLF), given their phase information. The PLF method is used in this work to extract the necessary features from the EEG to allow its classification.

Additionally, in order to maximize the information output from the BCI system, so as to have enough information to control a 2D cursor, six motor tasks were considered. In particular, motor tasks from the lower members were used, in an attempt to evaluate if they provide useful BCI control signals. Thus, a hierarchical Support Vector Machine (SVM) classifier is used to distinguish the various motor tasks, and its performance is assessed with the Leave-One-Out Cross Validation (LOOCV) method.

1.3 Thesis Outline

The text is organized as follows: Chapter 2 presents the basic concepts related to Brain-Computer Interfacing, describing its definition, structure and types of systems; Chapter 3 provides all the fundamental knowledge necessary to understand the physiologic origins of the EEG signals, including a brief description of the Nervous System, the brain waves measured by the EEG, and the effects of motor tasks on the dynamics of brain activity and its relevance in the context of a BCI system; Chapter 4 details all the methods used in the processing of the signals and their acquisition, with emphasis on the PLF method; Chapter 5 presents all the results obatined, which are discussed in Chapter 6, concluding the dissertation.

2

BCI Definition and Structure

Contents

In this chapter the definition of a BCI system is presented, including its structure and types of systems. Throughout the text, references to examples of actual BCI systems allow to illustrate the state of the art of this research field.

2.1 BCI Definition

Development of Brain-Computer Interface systems is one of the fastest growing areas in scientific research, evolving in the past 20 years, with more than 400 groups worldwide currently pursuing this field of investigation [4]. It encompasses various knowledge fields such as Medicine and Physiology, Instrumentation, Signal Acquisition and Processing and Computer Science, integrating these distinct areas into one single goal: extract and interpret information directly from the brain.

Until recently, BCI research has essentially followed a proof-of-concept strategy, attempting to demonstrate that useful information can indeed be extracted directly from the brain through a measure of its activity. The first experiments with BCIs were conducted by Schmidt *et al.* in the 1970's, who investigated whether primates could get control over the fire rate of individual neurons in the Primary Motor Cortex [5]. One of the first persons to benefit from BCI research was Matt Nagle, who was a C3 tetraplegic, paralyzed from the neck down. An electrode array was surgically implanted on the surface of his brain over a region of the motor cortex. The system was then trained to recognize the patient's thought patterns and translate them into commands to control a computer mouse [6]. Despite the good results provided by these early systems, they were very invasive, requiring surgery to place the electrodes on the relevant brain areas. Currently, researchers tend to use non-invasive methods, particularly through EEG measurements.

A BCI is defined as a system that measures and analyses brain signals and converts them in real time into outputs that do not depend on the normal output pathways of peripheral nerves and muscles [7]. It can be inferred from this definition that in order to have successful BCI operation, a closed loop of information is necessary between two adaptive controllers: the user, who produces specific brain signals that encode intent, and the BCI, which translates these signals into outputs that accomplish the user's intent, providing feedback to the user [4].

2.2 BCI Structure

A BCI, regardless of its purpose or recording methods, consists of four essential elements [4], as shown in Figure 2.1. These elements are described in detail bellow.

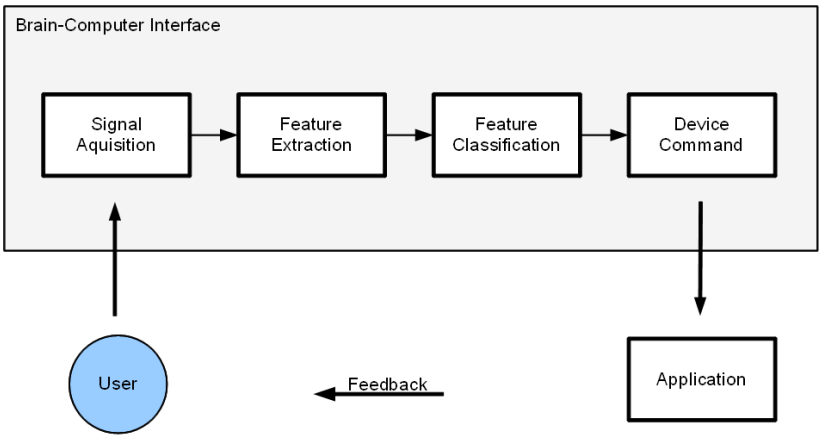


Figure 2.1: Basic BCI structure.

2.2.1 Signal Acquisition

The first step in BCI operation is the acquisition of brain signals. These brain signals include Electrophysiological, Magnetic and Metabolic signals. Electrophysiological systems measure the electrical activity of brain cells with electrodes either placed on the scalp (EEG), on the cortical surface (ECoG) or with intracortical devices. The Magnetic signals include essentially the Magnetoencephalogram (MEG), while the metabolic activity can be obtained from functional Magnetic Resonance Imaging (fMRI) or from Near-Infrared Spectroscopy (NIRS). Of these methods, the most widely used for BCI operation is the EEG, as it is a noninvasive, cost effective and simple technique, with good temporal resolution [4]. On the other hand, the EEG is more susceptible to noise and artifacts such as Electrooculogram (EOG), Electromyogram (EMG) or power line induced noise.

Independent of the recording method used, to operate a BCI users have to acquire conscious control over their brain activity [1]. Thera are two ways of accomplishing this (adapted from [1]):

1. Subjects percieve a set of stimuli dispayed by the BCI system and can control their brain activity by focusing onto one specific stimulus (Event-Related Potentials; e.g. P300, VEP, MRP);

2. Users control their brain activity by concentrating on a specific mental task (e.g. imagining motion of the right hand):

   - Oscillatory brain activity (sinusoid-like):
     - Delta band (1 - 4 Hz);
     - Theta band (4 - 8 Hz);
     - Alpha band (8 - 13 Hz);
     - Mu band (alpha activity over the sensorimotor cortex);
     - Beta band (13 - 25 Hz);
     - Gamma band (25 - 40 Hz);

- Slow Cortical Potentials (1 - 2 Hz);
- Neuronal Ensemble Activity (neuronal firing rate).

Note that current BCI systems do not read the user's mind and do not access private thoughts. Instead they rely on the natural physiologic response of the brain to a certain stimulation (whether internal or external), much like measuring an increase in heart rate when someone is excited or scared. Additionally, brain signal acquisition systems in use today do not have the capacity to measure or process each and every neuronal cell, much less every synapse established between neurons. Furthermore, there is not yet a widely accepted theoretical framework that explains how the brain works, although [8] may provide a good hint.

2.2.2 Feature Extraction

Having obtained the appropriate brain signals, the next step is feature extraction. Note that some systems imply some kind of signal preprocessing or enhancement before the feature extraction step in order to increase the signal-to-noise ratio, thus improving the overall quality of the system [9]. This is especially true in EEG-based systems given that, as mentioned before, these are more susceptible to noise. According to [9], the most common signal enhancing techniques used are Referencing Methods (such as the Common Average Reference (CAR) and the Surface Laplacian (SL)) and Common Spatial Patterns (CSP). Additionally to these methods, some research groups attempt to minimize the effects of EOG or EMG artifacts in the EEG signal. For example in [10] a linear combination and regression method is used to remove the EOG from the EEG. In [11] a combination of Blind Source Separation (BSS) and SVMs is used to isolate EOG and EMG artifacts into independent components from the EEG, and subsequent classification and removal of those components.

It is easy to understand that choosing the feature extraction method depends on the type of mental task on which the BCI is based. These methods can either generate features in the time domain or features in the frequency domain. Time domain feature extraction methods include filtering, wavelet transform and parametric models such as Auto-Regressive (AR) models. Frequency domain methods mainly rely on band power estimation using Welch's Method, Morlet Wavelets and AR models.

Additionally to the previous feature extraction methods, spatial domain features are also used in BCI signal processing. Examples of spatial domain features include Laplacian Filters, CSP, Independent Component Analysis (ICA) or simply the manual selection of the most relevant electrode channels for a given purpose. Note that spatial filters are an important aspect in BCI design since different mental tasks correspond to different brain areas. For instance, the Primary Motor Cortex (PMC), responsible for planning and executing movements, is located in the posterior portion of the frontal lobe. This cortical area can be subdivided into specific locations for the various limbs (see Figure 2.2). With this information it is possible to correlate brain activity at a specific area in the motor cortex with movement in a certain limb.

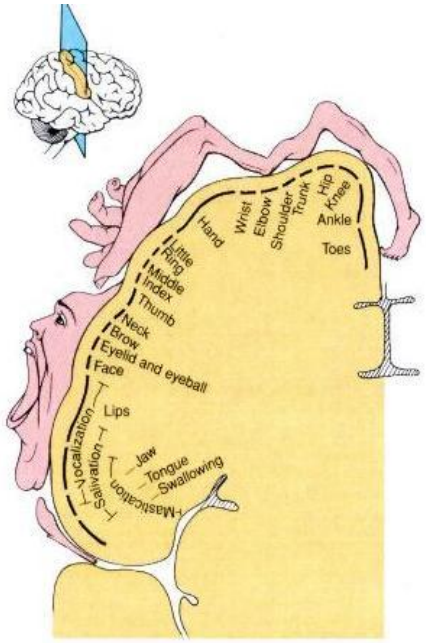


Figure 2.2: Mortor cortex in right cerebral hemisphere.

2.2.3 Feature Classification

Once the features have been obtained from the EEG, it is necessary to classify them according to the experiment (paradigm) characteristics. For example, distinguish the brain activity generated by moving the right hand from the brain activity arising from left hand movement. Most classification techniques rely on Machine Learning algorithms like Neural Network (NN), Hidden Markov Model (HMM), Threshold Detectors, Fisher's Discriminant Analysis (FDA), Support Vector Machine (SVM) and Linear Discriminant Analysis (LDA) [9]. Note that the feature space may be high-dimensional which makes necessary some kind of feature selection in order to maintain a computationally efficient system. On the other hand, the optimal sub-space is not known *a priori*, making the feature selection process difficult. In [12] a Genetic Algorithm (GA) is used to select feature sub-spaces, wrapped around a SVM classifier. The GA evolution is directed by the system's accuracy in identifying the mental task. This approach leads to higher classification rates than both the full feature set and randomly selected subsets [12]. Naturally, one can always manually choose feature sub-spaces and assess which sub-space produces higher classification rates.

2.2.4 Device Output

The first three components of a BCI system transform the user's intent, measured from brain activity, into class labels, with each label pertaining to a specific mental task as defined by the experimental paradigm. The next step encompasses the conversion of the class labels into actual commands that operate the interface.

Following the classification presented in [4], one can distinguish two types of BCI applications: 1) direct control of assistive technologies and 2) neurorehabilitation.

The most traditional BCI applications belong to the first group with examples raging from communication channels, movement and environment control to locomotion and even gaming and virtual environments. In [13] an assessment is made about the feasibility of walking through a virtual city by thought, by using motor imagery, combining spectral analysis as the feature extraction method with LDA as classifier. In [14] the EEG signal is used to control a prosthetic hand. The EEG is processed with a Laplacian Filter, from which an AR model is extracted. The features are classified with a hierarchical linear classifier, allowing the system to differentiate between three states: relaxation, and right versus left hand motor imagery (used to open and close a prosthetic hand). Additionally, local machine control, based on information from force and position sensors on th e hand, is used to achieve a stable grip. Another example of an assistive BCI application is presented in [15] which relies on a P300 based BCI to control the environment of a house.

The second group includes therapeutic applications, i.e. these applications serve as tools to help people whose neuromuscular function has been impaired by trauma or disease [4]. For instance [16] presents an overview of BCIs as a rehabilitation tool for tetraplegic patients. The BCI in this example aids the patient learn how to modify the efficacy of spared neuronal ensembles that represent movement, sensation and cognition through progressive practice with feedback and reward.

Although BCIs are no longer restricted to laboratories (the Wadsworth research group has been adapting their BCI system for home use [17]; California-based NeuroSky provides BCI systems for gaming applications [18]), much still needs to be done. Problems such as asynchronous BCI (a system that detects that the user is trying to communicate through the BCI) or subject and system training remain very much open [1].

3

Nervous System and Electrophysiology

Contents

This chapter deals with the basic physiology of the Nervous System, describing how signals are transmitted in the brain and how populations of neurons form networks that produce brain waves visible in EEG. Also, the fundamentals of ERD/ERS are presented and their relation to the brain waves is discussed. This is followed by a description of the neurophysiology of motor tasks, detailing which parts of the Nervous System are involved and which changes are induced in the EEG.

3.1 Nervous System

The Nervous System is a group of organs that transmit signals between different parts of the body. These signals attempt to control the various bodily activities. This is achieved by controlling the contraction of appropriate skeletal muscles, the contraction of the smooth muscle in the internal organs and the secretion of active chemical substances by both exocrine and endocrine glands.

The fundamental unit of the Nervous System is the neuron. Its most important property is the ability to communicate with other cells through synapses, which are membrane-to-membrane junctions that allow rapid transmission of signals either electrically or chemically. See Figure 3.1 for a depiction of a typical neuron.

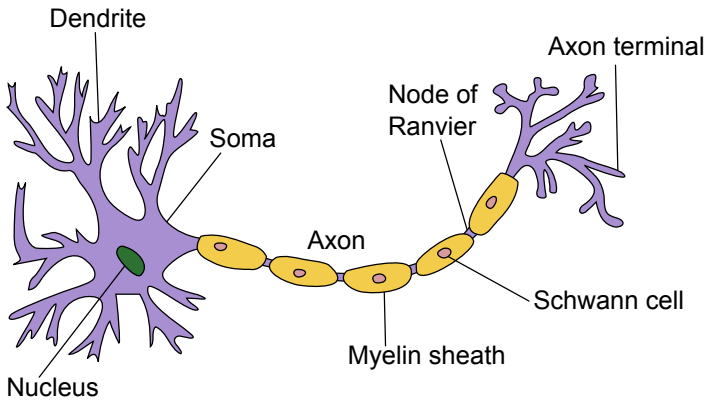


Figure 3.1: Structure of a typical neuron.

At a synapse, the plasma membrane of the presynaptic neuron comes into close apposition with the membrane of the postsynaptic cell. At these sites there are extensive arrays of membrane structures that allow the transmission of a signal. The characteristics of these structures depend on the type of synapse:

- Chemical Synapse — the presynaptic neuron releases a neurotransmitter that binds to receptors located in the postsynaptic cell;
- Electrical Synapse — the presynaptic and postsynaptic cell membranes are connected by channels that are capable of passing electrical current, causing voltage changes in the presynaptic cell to induce voltage changes in the postsynaptic cell.

The most frequent type of synapse found in the brain is the chemical synapse. This synapse, although not as fast as an electrical synapse, is strictly unidirectional and, depending on the type of neurotransmitter and postsynaptic cell, can either be excitatory or inhibitory. It also allows the modulation of the signal gain. Thus, the synapses perform a selective action, often blocking weak signals while allowing strong signals to pass, but at other times selecting and magnifying certain weak signals, and often channeling these signals in many directions rather than only one direction. Note that the estimated number of neurons in the human brain is in the order of 10^{11} and the number of synapses is in the order of 10^{14}. This makes the Nervous System a very complex network of cells constantly sending and receiving signals. How this network is organized is described in the following section.

3.1.1 Organization of the Nervous System

Anatomically, the Nervous System can be divided into two main subsystems: the Central Nervous System (CNS) and the Peripheral Nervous System (PNS). The CNS includes the brain and the spinal cord, while the PNS includes all the other nervous structures that do not reside in the CNS, feeding the external world to the CNS (through sensory, or afferent, pathways) and executing the commands dictated by the CNS (through motor, or efferent, pathways).

The PNS can be further subdivided into the Somatic Nervous System, which consists of nerves that innervate the skin, joints and muscles, and into the Autonomic Nervous System, whose neurons innervate the internal organs, blood vessels and glands. Thus, while the Somatic Nervous System is more involved with voluntary and conscious control of muscle activity, the Autonomic Nervous System is more involved in the subconscious maintenance of body functions, such as the heart rate, digestion, respiration rate, and others. Figure 3.2 summarizes these concepts.

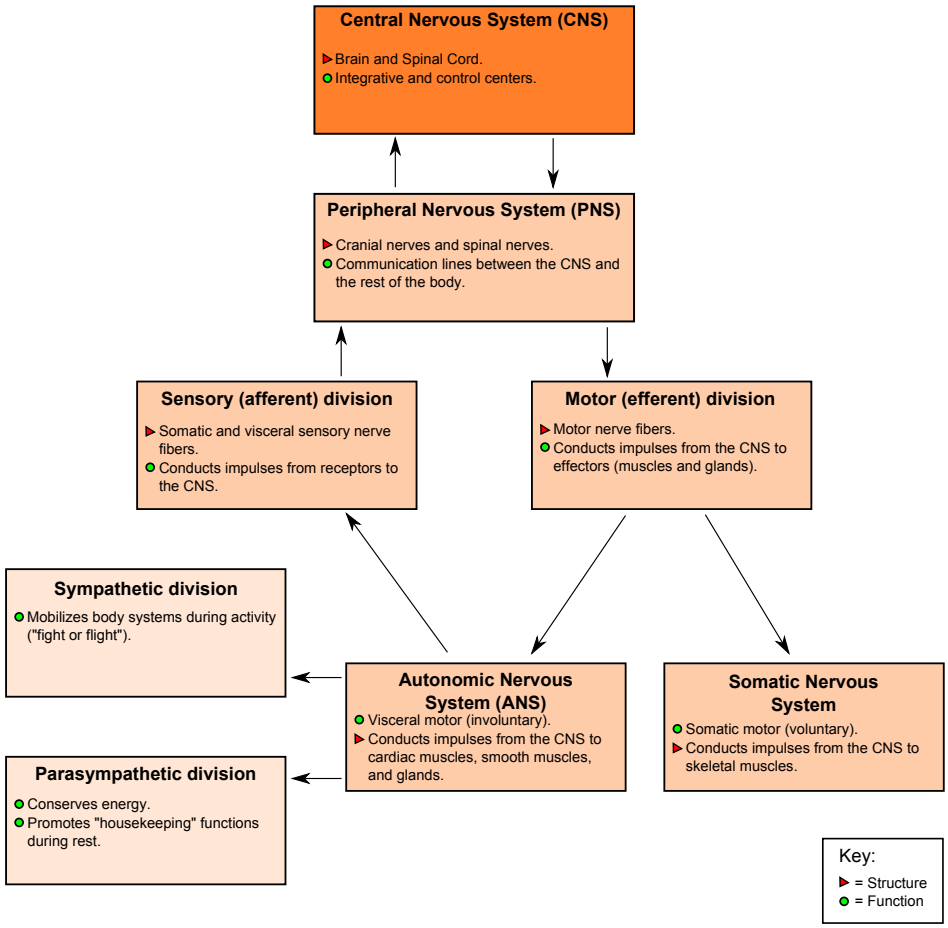


Figure 3.2: Organization of the Nervous System.

3.1.2 Major Levels of Central Nervous System Function

Given that a BCI takes its information from the CNS, and in particular the brain, it is relevant to understand how this part of the Nervous System is organized.

The human Nervous System has inherited special functional capabilities from each stage of human evolutionary development. From this heritage, three major levels of the central nervous system have specific functional characteristics: the spinal cord level, the lower brain or subcortical level, and the higher brain or cortical level.

3.1.2.A Spinal Cord Level

The spinal cord is frequently regarded as being only a conduit for signals from the periphery of the body to the brain, or in the opposite direction from the brain back to the body. This is far from the truth. For instance, neuronal circuits in the cord can cause walking movements, reflexes that withdraw portions of the body from painful objects, reflexes that stiffen the legs to support the body against gravity, and reflexes that control local blood vessels, gastrointestinal movements, or urinary excretion [19]. In fact, the upper levels of the nervous system often operate not by sending signals directly to the periphery of the body but by sending signals to the control centers of the cord, simply ”commanding” the cord centers to perform their functions.

3.1.2.B Lower Brain Level

The lower areas of the brain include the medulla, pons, mesencephalon, hypothalamus, thalamus, cerebellum, and basal ganglia. These areas control most of what are usually called subconscious activities of the body. For instance, subconscious control of arterial blood pressure and respiration is achieved mainly in the medulla and pons. Control of equilibrium is a combined function of the older portions of the cerebellum and the reticular substance of the medulla, pons and mesencephalon. And many emotional patterns, such as anger, excitement, sexual response, reaction to pain, and reaction to pleasure, can still occur after the destruction of much of the cerebral cortex.

3.1.2.C Cortical Brain Level

It seems, from the previous two levels, that there is nothing left for the cerebral cortex to do. This is not true, although a precise and detailed knowledge of its functions and how it works is not yet available.

The cerebral cortex is, in fact, a very large memory storehouse, never functioning alone but always in association with lower centers of the nervous system. Without the cerebral cortex, the functions of the lower brain centers are often imprecise, but the information stored in the cortex converts these functions to determinative and precise operations. Additionally, the cerebral cortex is essential for most of our thought processes, but it cannot function by itself. In fact, it is the lower brain centers, not the cortex, that initiate wakefulness in the cerebral cortex, thus opening its bank of memories to the thinking machinery of the brain [19].

3.1.3 Neuronal Networks

The brain processes information by organizing itself into populations of neuronal elements that form complex aggregates. These populations are generally called neuronal networks, or pools. One of the characteristics of such networks is that the underlying neuronal elements, under appropriate circumstances, tend to work in synchrony. Frequently, synchronous behavior is associated with oscillatory dynamics. Note that the discharge of a single excitatory presynaptic terminal almost never causes an action potential in a postsynaptic neuron. Instead, large numbers of input terminals must discharge on the same neuron either simultaneously or in rapid succession to cause excitation. We must also remember that some incoming fibers inhibit neurons, rather then exciting them. This emphasizes the importance of the network structure for the processing of information.

The CNS is composed of thousands to millions of neuronal networks; some of these contain few neurons, while others have vast numbers. Each neuronal network has its own special organization that causes it to process signals in its own unique way, thus allowing the total consortium of networks to achieve the multitude of functions of the nervous system. However, despite their different goals, the networks have similar principles of function.

Almost every part of the brain connects either directly or indirectly with every other part, and this can create a serious problem. If there were not mechanisms in the Nervous System to control the neuronal networks, the brain would be inundated by a mass of uncontrolled reverberating signals – signals that would be transmitting no information yet, nevertheless, would be consuming the circuits of the brain so that none of the informational signals could be transmitted. Such an effect occurs in widespread areas of the brain during epileptic seizures. Instead, in the healthy brain, two main mechanisms guarantee that feedback is only established in appropriate ways: inhibitory circuits, and fatigue of synapses.

There are two types of inhibitory circuits in the brain:

- Inhibitory feedback circuits that return from the termini of pathways back to the initial excitatory neurons of the same pathways — these circuits occur in virtually all sensory nervous pathways and inhibit either the input neurons or the intermediate neurons in the sensory pathway when the termini become overly excited;

- Gross inhibitory networks — these networks exert inhibitory control over widespread areas of the brain; for instance, many of the basal ganglia exert inhibitory influences throughout the muscle control system.

Synaptic fatigue means simply that synaptic transmission becomes progressively weaker the more prolonged the period of excitation. Furthermore, the shorter the interval between successive stimuli, the lower the intensity of the subsequent response.

3.1.4 Brain Waves

The discharge of a single neuron or single nerve fiber in the brain can never be recorded from the surface of the head. Instead, many thousands or even millions of neurons or fibers must fire synchronously. Only then will the potentials from the individual neurons or fibers summate enough to be recorded all the way through the skull. Thus, the intensity of the brain waves from the scalp is determined mainly by the number of neurons and fibers that fire in synchrony with one another, not by the total level of electrical activity in the brain. In fact, nonsynchronous nerve signals often nullify one another in the recorded brain waves because of opposing polarities. This implies that the EEG signals are a direct consequence of the existence of neuronal networks.

The character of the waves is dependent on the degree of activity in respective parts of the cerebral cortex, and the waves change markedly between the states of wakefulness, sleep, and coma. Much of the time, the brain waves are irregular, and no specific pattern can be discerned in the EEG. At other times, distinct patterns do appear, some of which are characteristic of specific abnormalities of the brain, such as epilepsy.

Brain waves are typically characterized by their frequency, with four main groups of waves:

- Delta Waves

  Delta waves include all the waves of the EEG with frequencies lower than 3.5 Hz, and they often have amplitudes two to four times greater than most other types of brain waves. They occur in very deep sleep, in infancy, and in some brain diseases.

- Theta Waves

  Theta waves have frequencies between 4 and 7 Hz. They occur normally in the parietal and temporal regions in children, but they also occur during emotional stress in some adults, particularly during disappointment and frustration. Theta waves also occur in many brain disorders, often in degenerative brain states.

- Alpha Waves

  Alpha waves are rhythmical waves that occur at frequencies between 8 and 13 Hz and are found in the EEGs of almost all normal adult people when they are awake and in a quiet, resting state of cerebration. These waves occur most intensely in the occipital region but can also be recorded from the parietal and frontal regions of the scalp. Their amplitude usually is about 50 μV. During sleep, the alpha waves disappear.

- Beta Waves

  When the awake person's attention is directed to some specific type of mental activity, the alpha waves are replaced by asynchronous, higher-frequency but lower-amplitude beta waves. These waves occur at frequencies greater than 14 Hz. They are recorded mainly from the parietal and frontal regions during specific activation of these parts of the brain.

It is known since Berger (1930) that certain events can block or desynchronize the ongoing alpha activity. These changes, which are time-locked, but not phase-locked, to the event, represent

frequency specific changes of the ongoing EEG activity and may consist, in general terms, either of decreases — termed Event-Related Desynchronization (ERD) — or increaes — termed Event-Related Synchronization (ERS) — of power in given frequency bands [20].

In general, the frequency of brain oscillations is negatively correlated with their amplitude, which means that the amplitude of fluctuations decreases with increasing frequency [20]. Also, there is a correlation between the level of cerebral activity and the average frequency of the EEG rhythm, with the average frequency increasing progressively with higher degrees of activity [19]. So, during periods of mental activity, the average frequency of the EEG increases, with the waves becoming asynchronous rather than synchronous, which leads to a decrease of the amplitude of the brain waves.

In summary, ERD/ERS reflect the dynamics between main neurons and interneurons that control the frequency components of the ongoing EEG. While ERD is correlated with activated cortical areas, ERS, characterized by synchronized alpha band rhythms during mental inactivity, introduces inhibitory effects. Mental tasks such as perceptual, judgment and memory tasks induce ERD/ERS. For instance, an increase in task complexity or attention results in an increase of the magnitude of ERD, in particular in the lower alpha band (7 – 10 Hz), being topographically widespread. On the other hand, voluntary movements result in a circumscribed ERD in the alpha band.

3.2 Neurophysiology of Motor Tasks

Most voluntary movements initiated by the cerebral cortex are achieved when the cortex activates patterns of function stored in lower brain areas: the cord, brain stem, basal ganglia, and cerebellum. These lower centers, in turn, send specific control signals to the muscles. For a few type of movements, however, the cortex has almost a direct pathway to the anterior motor neurons of the cord, bypassing some motor centers on the way. This is especially true for control of the fine dexterous movements of the fingers and hands.

3.2.1 The Motor Cortex

Figure 3.3 shows the functional areas of the cerebral motor cortex. The motor cortex itself is divided into three subareas, each of which has its own topographical representation of muscles groups and specific motor functions: the PMC, the premotor area, and the supplementary motor area.

Primary Motor Cortex

The PMC, shown in Figure 3.3, lies in the first convolution of the frontal lobe, anterior to the central sulcus. The figure lists the approximate topographical representations of the different muscle areas of the body in the PMC. This topographical organization is demonstrated even more graphically in Figure 3.4, which shows the degrees of representation of the different muscle areas as mapped by Penfield and Rasmussen. Note that due to the crossing of the nervous pathways in spinal cord, muscles of the right-hand side of the body are represented on the left-hand part of the sensorimotor cortex and

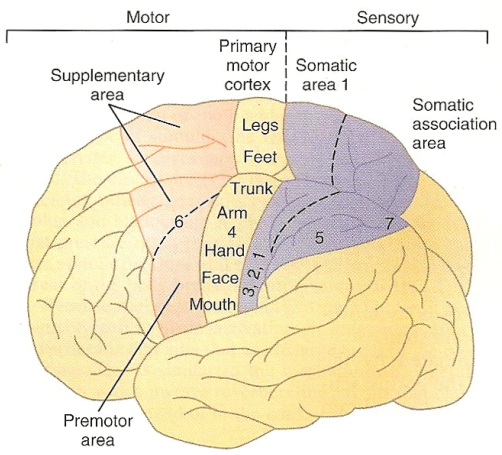


Figure 3.3: Motor and somatosensory functional areas of the cerebral cortex [19].

vice-versa. The mapping of the motor cortex was done by electrically stimulating the different areas of the motor cortex in human beings who were undergoing neurosurgery. Note that more than one half of the entire PMC is concerned with controlling the muscles of the hand and the muscles of speech. Point stimulation in these hand and speech areas on rare occasions causes contraction of a single muscle. Most often, stimulation contracts a group of muscles instead. To express this in another way, excitation of a single motor cortex neuron usually excites a specific movement rather than one specific muscle.

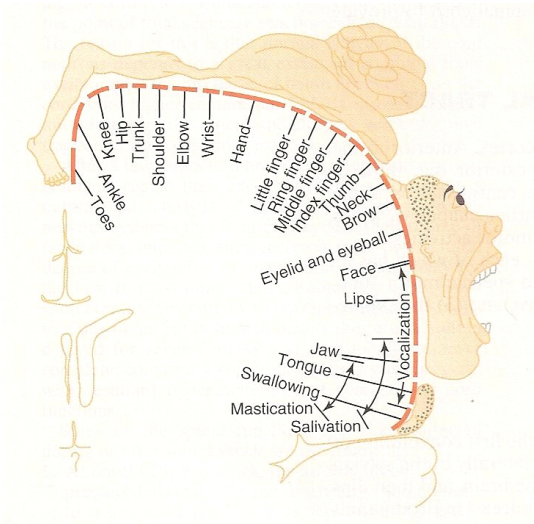


Figure 3.4: Degree of representation of the different muscles of the body in the motor cortex [19].

Premotor Area

The premotor area lies 1 to 3 cm anterior to the PMC. The topographical organization of the premotor cortex is roughly the same as that of the PMC. Nerve signals generated in the premotor area cause much more complex patterns of movement that the discrete patterns generated in the PMC. For instance, the pattern may be to position the shoulders and arms so that the hands are properly

oriented to perform specific tasks. To achieve these results, the most anterior part of the premotor area first develops a ”motor image” of the total muscle movement that is to be performed. Then, in the posterior premotor cortex, this image excites each successive pattern of muscle activity required to achieve the image. This posterior part of the premotor cortex sends its signals either directly to the PMC to excite specific muscles or, often, by way of the basal ganglia and thalamus back to the PMC. Thus, the premotor cortex, basal ganglia, thalamus, and primary motor cortex constitute a complex overall system for the control of complex patterns of coordinated muscle activity.

Supplementary Motor Area

The supplementary motor area has yet another topographical organization for the control of motor functions. Contractions elicited by stimulating this area are often bilateral rather than unilateral. For instance, stimulation frequently leads to bilateral grasping movements of both hands simultaneously. In general, this area functions in concert with the premotor area to provide body-wide attitudinal movements, fixation movements of the different segments of the body, positional movements of the head and eyes, and so forth, as background for the finer motor control of the arms and hands by the premotor area and the PMC.

Motor signals are transmitted directly from the cortex to the spinal cord through the corticospinal tract and indirectly through multiple accessory pathways that involve the basal ganglia, cerebellum, and various nuclei of the brain stem. In general, the direct pathways are concerned more with discrete and detailed movements, especially of the distal segments of the limbs, particularly the hands and fingers. When nerve signals from the motor cortex cause a muscle to contract, somatosensory signals return all the way from the activated region of the body to the neurons in the motor cortex that initiated the action. These somatic signals often cause positive feedback enhancement of the muscle contraction.

3.2.2 ERD/ERS in Movement Tasks

Movement tasks induce changes to the ongoing dynamics of mental activity. Of these changes, ERD/ERS assumes particular relevance in the context of BCI. Following the work done by Pfurtscheller and Lopes da Silva in [20], a detailed description of what happens in the brain, in respect to the ERD and ERS dynamics during a movement task, is now presented.

Firstly, it has to be stressed that ERD/ERS is not a unitary phenomenon. Rather, it is comprised of various separate events, occurring at different frequency bands, with distinct starting instants and durations. The most dominant feature is a desynchronization in the upper alpha band (10 – 12 Hz) and in the lower beta band (13 – 20 Hz), which is topographically restricted to the corresponding sensorimotor region (contralateral to the movement). It starts up to 2 s prior to movement onset, and becomes bilaterally symmetrical immediately before execution of movement. Furthermore, ERD is accompanied by ERS over unrelated cortical areas and can be accompanied by ipsilateral ERS (on the same side of the movement), circumscribed and localized, in the beta band, to the correspond-

ing sensorimotor region. Also, there is a post-movement beta ERS found in the first second after termination of movement, which is focused around the corresponding sensorimotor area and whose maximum coincides with a reduced excitability of motor cortex neurons.

Note that these events were described for the actual performance of movement tasks. However, it has been shown that motor imagery and motor observation activate similar neuronal networks to those activated by actual movement [21]. Thus, it is expected that the imagination of motor tasks, that is to say, the conscious effort of simulating movements without actually performing them in an overt fashion [21], leads to similar phenomena in the brain as a real movement, such as the ERD/ERS dynamics just described. This is of particular importance in the context of a BCI, as its definition implies that the system does not depend on the normal output pathways of peripheral nerves and muscles. Accordingly, the imagination of motor tasks, while producing similar results as actual movement, allows to comply with the definition of a BCI, providing a means to control it, with virtually all of the movements available to the human being [1]. Furthermore, a BCI system based on the imagination of motor tasks might be more intuitive in some types of applications of the kind of navigating through a virtual city by thought, as is done in [13], as the cortical pathways previously used to produce movements continue to do so through the BCI.

[1] Naturally, there are movements that are easier to detect than others.

4

Methods

Contents

This chapter presents all the methods employed for the development of the BCI system. The chapter is broken into four sections. The first section describes the experimental setup of the EEG acquisitions, detailing the system configuration and the experimental procedure. The second section presents the preprocessing steps employed to prepare the acquired signal for the main processing stage, which is portrayed in the third section. The fourth and last section details the approach taken to classify the various motor tasks. All the processing steps were implemented in MATLAB®.

4.1 Experimental Setup

The acquired signals consist of EEG data from 6 subjects (2 female, 4 male, ages (22.3 ± 0.5) years, all right handed). The subjects were comfortably sitting on a chair in front of a CRT computer screen, which conducted the subjects throughout the experiment. The subjects were fitted with a 64-electrodes cap (following the standard $10 - 20$ system – see Figure 4.1) connected to a Brain Products' QuickAmp amplifier. The EEG channels were recorded with a sampling rate of $2000\ Hz$, and were stored in a separate computer from the one conducting the experiment. Synchronization of the recording with the various stimuli was done by connecting the first computer to the amplifier via a parallel cable.

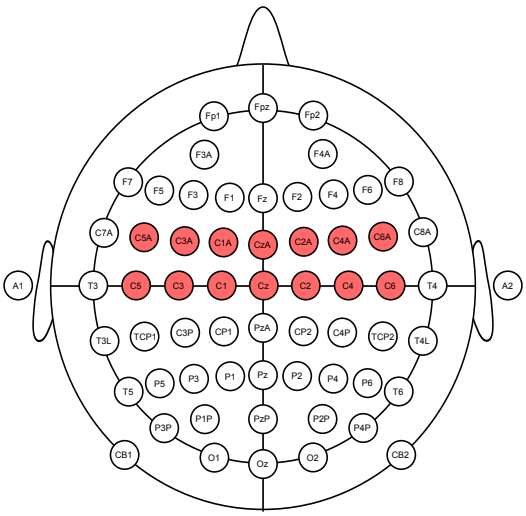


Figure 4.1: Position of the EEG electrodes; red electrodes are placed over the Primary Motor Cortex.

The experiment was conducted with the E-Prime 2.0 software. The cue-based BCI paradigm consisted of seven different motor tasks:

- Class 1 - no movement (CC);
- Class 2 - movement of the right foot (RF);
- Class 3 - movement of the left foot (LF);

- Class 4 - movement of the right leg (RL);
- Class 5 - movement of the left leg (LL);
- Class 6 - movement of the right hand (RH);
- Class 7 - movement of the left hand (LH).

One session was recorded with each subject. The sessions comprised a brief introduction to the experiment followed by two runs separated by a short break. Each run consisted of two groups of trials, with the first group dedicated to actual realization of the motor tasks, while in the second group users were asked to imagine the motor tasks. Each group is comprised of three cycles through the motor tasks, with the sequence of the tasks presented in one cycle exhibiting a random order. Each trial starts with the presentation of a fixation cross over a blank screen. After $1000\ ms$ a figure appears indicating the motor task to be executed, lasting for $4000\ ms$. At the end of this period both the fixation cross and figure are replaced with a relaxation indication, giving the subjects the opportunity to blink, lasting for $2000\ ms$. A final blank screen ($1000\ ms$) allows the transition to the next trial. See Figure 4.2 for a graphical representation of the trial structure.

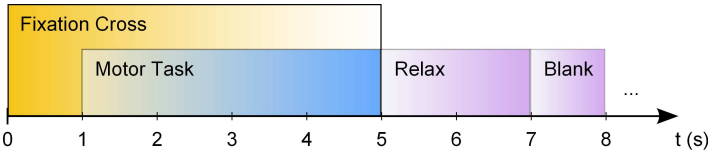


Figure 4.2: Structure of a trial of the EEG recording sessions.

The acquisitions were exported from the amplifier software (BrainVision Analyzer) in the form of text files, in order to be processed later on.

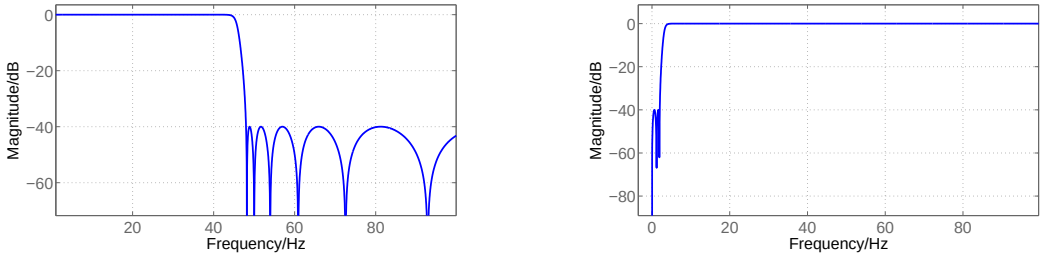
4.2 Signal Preprocessing

The files generated for each subject are of the size of, approximately, $1\ GB$. This poses some practical difficulties in the processing of the signals, particularly in systems with restricted memory. Additionally, current BCI researchers do not use high sampling rates. Therefore, a downsampling step was used in the preprocessing stage to reduce the sampling rate from $2000\ Hz$ to $500\ Hz$. This was done by taking every fourth sample from the original files.

Afterward, the signals were processed with two Chebyshev type II filters in order to remove the components below $5\ Hz$ and above $45\ Hz$. This type of filters, besides having a steeper roll-off than Butterworth filters, also has no ripples in the passband, but does have equiripple in the stop band. The order of the filters was chosen so that the magnitude in the stopband is maintained below $-40\ dB$. These frequencies were chosen because it is this frequency band that covers the brain activity of interest in this work (the alpha and beta bands)[1]. The signals were filtered with forward and reverse

[1]The passband is actually wider than the alpha and beta bands to avoid the effects of roll-off.

runs to obtain a zero phase distortion output. The magnitude responses of the filters are shown in Figure 4.3.



(a) Lowpass filter; $f_c = 45\,Hz$.

(b) Highpass filter; $f_c = 5\,Hz$.

Figure 4.3: Magnitude response of the two Chebyshev type II filters used in the preprocessing stage; f_c is the cutoff frequency.

Subsequently, the trials were isolated and ordered, allowing for a more efficient processing of the data. Finally, the signals were parsed with a small Laplacian Filter. A Laplacian Filter changes the reference of a given electrode to the mean of its nearest neighbors, as shown in Equation 4.1, where $x[n]$ is the original signal and S_x is the set of the C nearest channels to x. This enhances the discrimination of local activity.

$$x_{lap}[n] = x[n] - \frac{1}{C}\sum_{i \in S_x} x_i[n] \tag{4.1}$$

In this work, the four nearest neighbors of a given electrode (to the north, south, east and west) were used for the Laplacian Filter, which was applied to a subset of the original set of channels, chosen over the PMC, as shown in Figure 4.1.

4.3 Feature Extraction

As discussed previously, Event-Related Desynchronization and Event-Related Synchronization are the fundamental phenomena to be detected in the EEG in order to implement a BCI system based on motor tasks. It will now be described what measures (features) can be extracted from the EEG to detect ERD/ERS. The focus is given to band power methods and to an alternative method based on synchronization quantification, through which the synchronization between signals is measured.

4.3.1 Band Power

The traditional method to identify and measure ERD/ERS is by computing the power of the input signals in specific frequency bands. Indeed, about one-third of BCI designs have used power-spectral features [9]. To do so, there are several different techniques currently used in the development of BCI systems.

4.3.1.A Classical Method

The classical method, as described by Pfurtscheller and Lopes da Silva in [20], computes the time course of ERD/ERS as the EEG power within identified frequency bands, relative (as percentage) to the power of the reference or baseline period a few seconds before the event occurs. It includes the following steps:

1. Bandpass filtering of all event-related trials;
2. Squaring of the amplitude samples to obtain power samples;
3. Averaging of power samples across all trials;
4. Averaging over time samples to smooth the data and reduce the variability.

To obtain percentage values for the ERD/ERS power estimation, Equation 4.2 is used, where A is the power in the motor task (active) period and R is the power of the reference or baseline period.

$$Pwr\% = \frac{(A - R)}{R} \times 100 \tag{4.2}$$

This method, while appropriate to make a statistically significant assessment of ERD/ERS, is not adequate for use in a BCI system. This is so because the third step in the method averages the power samples across all trials, not allowing for a real time system, as mandated by the BCI definition. Furthermore, the procedure of the EEG acquisitions made for this dissertation only establish a time period in which the motor task is to be executed, which may not coincide with the actual period that the motor task is performed by the subject. This is due to different response times to the starting cue, different interpretations of the motor task to be done, subject concentration or fatigue, among others. Thus, the averaging step, adding to the fact that there are only six trials for each task, could potentially result in poor results.

Another disadvantage of this method is the fact that it is necessary to choose a frequency band, which is subject specif, *a priori*. Therefore, an additional step must be taken before any measures of ERS/ERS are made, such as the detection of the most reactive frequency band based on the analysis of short-term power spectra. This is easily done using the Fourier Transform (FT).

4.3.1.B Fourier Transform

The Fourier Transform (FT) of a signal $x(t)$, as defined by its analysis and synthesis expression in Equations 4.3 and 4.4, respectively, is a very useful mathematical tool to study the frequency structure of the signal.

$$X(\omega) = \int_{-\infty}^{\infty} x(t)\, e^{-j\omega t} dt \tag{4.3}$$

$$x(t) = \frac{1}{2\,\pi} \int_{-\infty}^{\infty} X(\omega)\, e^{j\omega t} d\omega \tag{4.4}$$

The use of the FT has provided many theoretical and practical results in the field of Signal Processing, with numerous applications. Naturally, for use in digital signals, it is necessary to adapt the definition of the FT to accommodate this type of signals, resulting in the Discrete Fourier Transform (DFT).

The DFT is defined in Equation 4.5 and its inverse is presented in Equation 4.6, for a discrete signal $x[n]$ with $n = 0, ..., N-1$.

$$X[\omega_k] = \sum_{n=0}^{N-1} x[n]\, e^{-j\frac{2\pi kn}{N}},\ k = 0, 1, ..., N-1 \tag{4.5}$$

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[\omega_k]\, e^{j\frac{2\pi kn}{N}},\ n = 0, 1, ..., N-1 \tag{4.6}$$

In summary, the DFT is proportional to the set of coefficients of projection onto the sinusoidal basis set, and the inverse DFT is the reconstruction of the original signal as a superposition of its sinusoidal projections.

From the definition of the DFT it is evident that $X[\omega_k]$ is only defined for frequencies $\omega_k = 2\pi k f_S/N$, where f_S is the sampling rate of the signal. If the signal is constituted by exact multiples of these frequencies, then the spectrum is zero everywhere except at $\omega = \omega_k, k \in [0, N-1]$. However, this is generally not true, and there is a spectral leakage effect, whereby the energy of a given frequency component is smeared across the entire spectrum, with a more evident peak at the closest ω_k. The simplest solution to avoid leakage would be to choose N to be exactly a period, although this is generally not practical. Usually, to reduce spectral leakage, window functions are used, such as the raised cosine window that tapers the data smoothly to zero at both endpoints of the window.

There is also the question of spectral resolution, which is determined by N. Evidently, it is not practical to arbitrarily increase N, as it would be necessary to acquire many data samples before any processing can be done. This is especially true for real-time processing, or if a short-time Fourier analysis is required. One trick frequently used to overcome this problem is zero padding the signal, i.e., adding a number of zeros to the signal. This technique, while not altering the information contained by the signal, nor adding any new information, provides a way to increase the frequency resolution by interpolating the available information in a higher number of frequency points. Additionally, zero padding can be used to increase the number of data points to a power of 2, which is computationally more efficient in some algorithms used to compute the DFT, such as the Fast Fourier Transform (FFT) algorithm, used in this work.

4.3.1.C Auto-Regressive Models

Auto-Regressive (AR) modeling utilizes the time history of a signal to extract important information hidden in the signal [22]. The AR model predicts the current values of a time series based on the hypothesis that the most recent data points contain more information than the other data points and, thus, each value of the series can be computed as a weighted sum of the previous values of the series

plus an error term, as can be inferred from Equation 4.7, where $x[n]$ is the current value of the time series, $a_1, ..., a_M$ are the predictor coefficients, M is the model order and $\varepsilon[n]$ represents a one-step prediction error.

$$x[n] = \sum_{k=1}^{M} a_k \, x[n-k] + \varepsilon[n] \tag{4.7}$$

Taking the z-transform from Equation 4.7, the following is obtained:

$$X(z) = \sum_{k=1}^{M} a_k \, X(z) \, z^{-k} + E(z) \tag{4.8}$$

This ca be rewritten as

$$H(z) = \frac{X(z)}{W(z)} = \frac{1}{1 - \sum_{k=1}^{M} a_k \, z^{-k}} \tag{4.9}$$

where $H(z)$ is the transfer function of the AR synthesis filter. $X(z)$ may be regarded as the output of the AR filter fed by the prediction error sequence $W(z)$. If z is substituted by $e^{j\omega}$, the spectrum of the modeled time series, $R(e^{j\omega})$, is obtained by multiplying the prediction error variance (σ_p^2) with the square of the transfer function, as shown in Equation 4.10.

$$R(e^{j\omega}) = \left|H(e^{j\omega})\right|^2 \, \sigma_p^2 = \frac{\sigma_p^2}{\left|1 - \sum_{k=1}^{M} a_k \, e^{-j\omega k}\right|^2} \tag{4.10}$$

With this formulation it is possible to evaluate the spectrum of the signal in an arbitrary number of frequency points, whereas with the DFT the frequency resolution if determined by the number of data points used for calculation. The AR model also enables the frequency components to be determined more exactly than with the DFT, provided that the model is good enough [22]. This implies that, in order for the model to represent well the signal, its order M must be adequate.

AR models have been used with success in the design of BCI systems, such as in [14], but there seems to be no concrete knowledge on the ”correct” model order to use in a BCI. In fact, the AR model order that optimizes a motor task BCI is generally higher than the model orders that are frequently used in BCI studies [23]. There are some objective methods to select the model order, such as the Akaike’s Information Criterion, which is a weighted combination of prediction error and model complexity, but it would seem that the correct approach, based on results obtained by McFarland and Wolpaw in [24], is that model order selection should be based on criteria that reflect system performance. This introduces one more variable that needs to be estimated by the system.

4.3.1.D Implementation

The band power features were obtained from the preprocessed EEG signals by using the DFT, implemented in MATLAB with the FFT algorithm, in successive Hamming windows of $256\,ms$ (128 samples) with 50% overlap. In each window, the power spectrum is computed with the FFT, with zero-padding to amount to 4096 frequency points. The average power in the frequency band between $8\,Hz$

and $15\,Hz$ is then computed. Therefore, for each window, a feature vector with 14 entries (the number of preprocessed EEG channels) is generated, with each entry representing the average power for that channel in the current window. Considering a trial with a duration of $8\,s$, which amounts to 61 windows, the results of the previous steps are stored in a matrix with dimensions 14×61. Each line of this matrix is then smoothed with an order 5 median filter [2]. A diagram of the procedure can be seen in Figure 4.4.

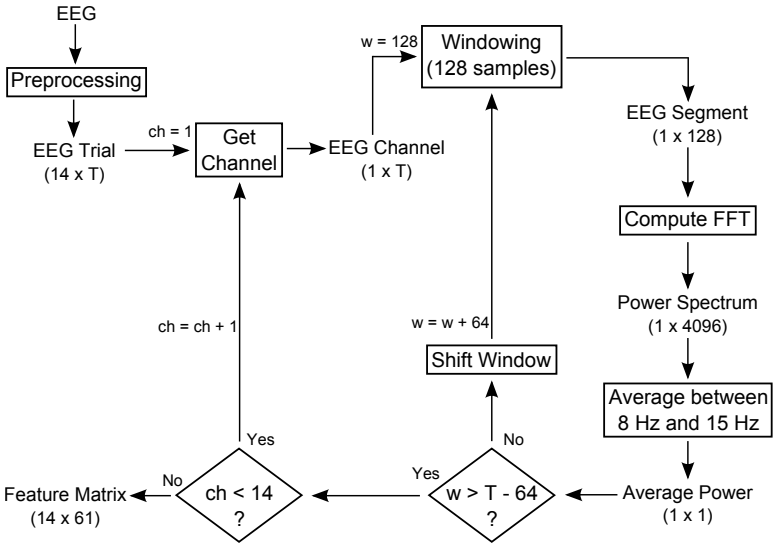


Figure 4.4: Diagram of the processing steps used to obtain the band power features; T $=$ 4000 samples is the trial length.

4.3.2 Phase-Locking Factor

As was discussed in Section 3.1, ERD arising from a motor task results from the loss of synchronization that specific neuronal networks, involved in the production of movement, experience during activity. This provides a framework on which to base the development of alternative methods of identifying ERD/ERS in the EEG. The main idea is that the analysis of the signals' phase could provide additional information to locate and identify motor tasks in the EEG.

This is not the first attempt at incorporating phase information into the analysis of mental activity. For instance, Kalitzin *et al.* use phase information in [25], as measured by the Phase Clustering Index (PCI), to predict the occurrence of epileptic seizures in patients with sensitivity to intermittent photic stimulation.

For motor task BCIs there is a group of features called Complex Band Power (CBP) features, as described in [26]. CBP features consist of two components: phase features and amplitude features. To extract the CBP features from a given signal, the DFT is computed, resulting in a set of complex coefficient representing equally spaced discrete frequency samples. Thus amplitude and phase of the original signal are obtained by simply computing the amplitude and angle of the complex coefficients.

[2] The median filter, while removing noise from a signal, preserves the edges of the objects contained in the signal (in this case, the ERD events); it is a lowpass filter.

Note that the absolute phase without respect to some reference is not meaningful [26]. Instead, it is how quickly the phase angle is changing, and in which direction, that is used in the CBP features.

A similar approach to the CBP features is used in the definition of the Phase-Locking Factor (PLF). The PLF is a method through which the synchronization between two signals is measured, provided there is access to their phase information. Given two oscillators with phases $\phi_i[n]$ and $\phi_k[n]$, $n = 1, ..., N$, the PLF is defined as [3]:

$$\varrho_{ik} = \left| \frac{1}{N} \sum_{n=1}^{N} e^{j(\phi_i[n] - \phi_k[n])} \right| \tag{4.11}$$

This measure ranges from 0 to 1. While the value $\varrho_{ik} = 1$ corresponds to perfect synchronization between the two signals, i.e., there is a constant phase lag, the value $\varrho_{ik} = 0$ corresponds to no synchronization, i.e., the phases are not correlated. Put simply, the PLF assesses whether the difference between the phases of the oscillators are strongly or weakly clustered around some angle in the complex unitary circle. Also, note that $\varrho_{ik} = \varrho_{ki}$.

In this work, the phase information is extracted from the EEG signals through the concept of Analytical Signals.

4.3.2.A Analytical Signals

An analytical signal, in the signal processing domain [3], is a signal which has no negative-frequency components. Therefore, in continuous time, every analytical signal $z(t)$ can be represented as:

$$z(t) = \frac{1}{2\pi} \int_0^{\infty} Z(\omega) e^{j\omega t} d\omega \tag{4.12}$$

where $Z(\omega)$ is the complex coefficient (setting the amplitude and phase) of the positive complex sinusoid $e^{j\omega t}$ at frequency ω.

For any real sinusoid $A\,cos(\omega t + \phi)$, a positive-frequency complex sinusoid $A\,e^{j(\omega t + \phi)}$ can be obtained by simply adding a phase-quadrature component $A\,sin(\omega t + \phi)$ as the imaginary part:

$$A\,e^{j(\omega t + \phi)} = A\,cos(\omega t + \phi) + jA\,sin(\omega t + \phi) \tag{4.13}$$

The relation expressed in Equation 4.13 implies that a positive-frequency analytical signal may be generated from the in-phase component by adding a quarter-cycle time shift. This notion extends to include any real signal, as they can be decomposed in a sum of sinusoids, by defining a filter that shifts each component by a quarter cycle. This is called the Hilbert Transform Filter. Ideally, this filter has unitary magnitude at all frequencies and introduces a phase shift of $-\pi/2$ at each positive frequency and $+\pi/2$ at each negative frequency. Given a real signal $x(t)$ and its Hilbert transform, denoted as $\mathcal{H}_t\{x\}$ for time instant t, the corresponding analytical signal $z(t)$ is obtained as:

$$z(t) = x(t) + j\,\mathcal{H}_t\{x\} \tag{4.14}$$

[3] In Mathematics, particularly in Complex Variables, a function is analytical if it satisfies certain properties regarding differentiability; the concept of analytical signals should not be confused with analytic functions.

The complex signal $z(t)$ thus defined has the property that all negative frequencies of $x(t)$ have been removed. To see how this works, consider the signal $x(t)$ defined as

$$x(t) = A\,cos(\omega_0 t) = \frac{A}{2}e^{j\omega_0 t} + \frac{A}{2}e^{-j\omega_0 t} \tag{4.15}$$

The Hilbert transform of $x(t)$ is given by:

$$y(t) = \mathcal{H}_t\{x\} = \frac{A}{2}e^{j\omega_0 t - j\pi/2} + \frac{A}{2}e^{-j\omega_0 t + j\pi/2} = -\frac{j\,A}{2}e^{j\omega_0 t} + \frac{j\,A}{2}e^{-j\omega_0 t} \tag{4.16}$$

By adding the signals from Equations 4.15 and 4.16 according to Equation 4.14, the analytical signal of $x(t)$ is obtained:

$$\begin{aligned} z(t) = x(t) + j\,y(t) &= \frac{A}{2}e^{j\omega_0 t} + \frac{A}{2}e^{-j\omega_0 t} + j\left(-\frac{j\,A}{2}e^{j\omega_0 t} + \frac{j\,A}{2}e^{-j\omega_0 t}\right) \\ &= \frac{A}{2}e^{j\omega_0 t} + \frac{A}{2}e^{-j\omega_0 t} + \frac{A}{2}e^{j\omega_0 t} - \frac{A}{2}e^{-j\omega_0 t} \\ &= A\,e^{j\omega_0 t} \end{aligned} \tag{4.17}$$

Note that the terms pertaining to negative-frequency components cancel each other, and there is also a gain of 2 at positive frequencies.This is true for any real signal. Another relevant observation is that there is no loss of information in the transformation of a signal into its analytical signal. This is due to the hermitian symmetry of the spectrum of a real signal, i.e., $X(-\omega) = X^*(\omega)$, where $(.)^*$ is the complex conjugate.

The use of analytical signals facilitates many mathematical manipulations, making certain attributes of the input signal more accessible, such as derivation of modulation and demodulation techniques. In particular, it allows easy access to both the amplitude envelope $A(t)$ and the instantaneous phase $\phi(t)$ of the input signal. These are obtained by simply taking the amplitude and angle of the analytical signal, as expressed in Equations 4.18.

$$\begin{aligned} A(t) &= |z(t)| \\ \phi(t) &= arg\,(z(t)) \end{aligned} \tag{4.18}$$

The concept of analytical signals thus provides a mean to obtain the phase information necessary to use the Phase-Locking Factor.

4.3.2.B Implementation

A similar approach to the band power features was taken here for the PLF features. As the PLF is a measure between two signals, a set of pairs needed to be defined. Therefore, from the 14 preprocessed EEG channels, 37 pairs were defined, as shown in Figure 4.5.

For each pair, the same windowing scheme as in the band power features was used (hamming windows of $256\,ms$ with 50% overlap). In each window, the analytical signal is obtained for both signals

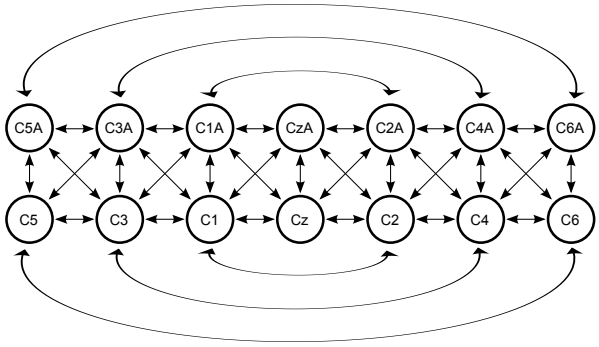


Figure 4.5: Pairs of EEG channels used to compute the PLF measure.

of the pair and their phases are extracted. The PLF is then computed with this information. Thus, a PLF value is obtained for each window of a given pair, resulting in a final matrix with dimensions 37×61. Each line of this matrix, which is also smoothed with an order 5 median filter, represents the evolution over time of the PLF for a certain pair of channels. A diagram of this procedure can be seen in Figure 4.6.

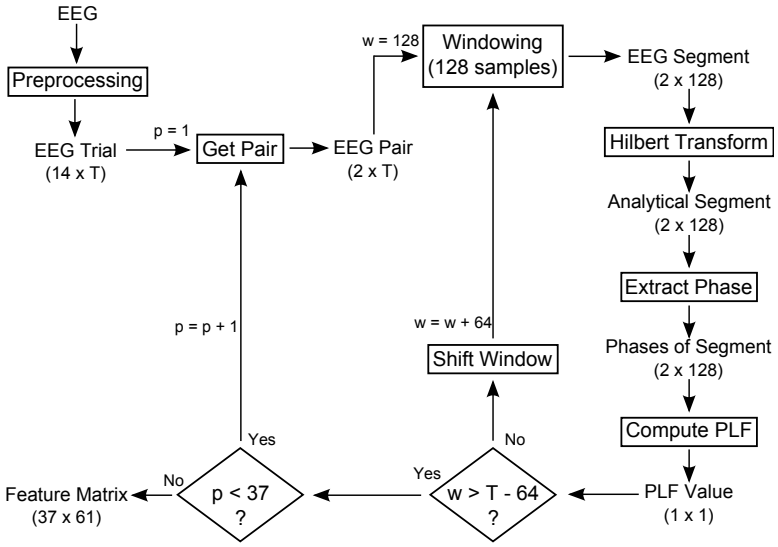


Figure 4.6: Diagram of the processing steps used to obtain the PLF features; T $=$ 4000 samples is the trial length.

4.4 Classification

Having described the feature extraction methods in the previous sections, it will now be presented the classification procedure used to identify the various motor tasks from the features. This encompasses the identification of ERD events in the time-evolution of the features and then feeding this information to a classifier, deciding which of the selected features belong to a certain motor task. These are all techniques in the Machine Learning field.

The field of Machine Learning is concerned with the question of how to construct computer programs that automatically improve with experience [27]. In a general way, in a machine learning problem, an input signal is processed with a given system, resulting in an output that provides the intended relevant information about the input signal. The task of automatically adapting the parameters and/or structure of the system, using a set of training data, constitutes the machine learning problem. There are two main approaches to this question: the supervised and the unsupervised routes. In the supervised approach, it is known *a priori* which elements of the training set belong to a certain class or exhibit a desired property. This is the case, for instance, of a classifier, where a pre-classified set of examples is used to train the system to classify new data. On the other hand, in an unsupervised approach, no information about the training set is available *a priori*, and so the system is left to adapt itself to the intrinsic properties of the data. Clustering, i.e., the grouping of data by similarity, is a typical example of an unsupervised approach.

In the following sections, the identification of ERD events is dealt with an unsupervised method, while the classification step is approached in a supervised way, using SVMs.

4.4.1 Identifying ERD

The identification of ERD in a motor task trial is done using a thresholding technique. The method employed only takes into account some of the lines from the feature matrix, i.e., the time-courses of two features. For the band power features, the power time-course of the channels C3, C1, Cz, C2, and C4 are used, whereas for the PLF features, the PLF time-courses C3-C1 and C4-C6 are taken into account.

Consider now that the selected time-courses are denoted by $f_i[n]$, with $n = 1, ..., N$, where N is the number of time samples. Before thresholding, the selected signals are fitted to order 3 polynomials, which are then subtracted from the original signals, according to Equation 4.19, generating the new signals $f_i^{new}[n]$. This allows for a more robust signal to be used in the thresholding step, reducing baseline wander. In other words, it is a highpass filter.

$$f_i^{new}[n] = f_i[n] - polyfit3(f_i, n) \tag{4.19}$$

The threshold, as defined in equation 4.20 with the mean and standard deviation, uses the first

second of the $f_i^{new}[n]$ signals as reference, as it is this time period that can be associated with a stationary state with higher certainty.

$$\begin{aligned} \mu_i &= \frac{1}{N_1}\sum_{n=1}^{N_1} f_i^{new}[n] \\ \sigma_i^2 &= \frac{1}{N_1-1}\sum_{n=1}^{N_1} (f_i^{new}[n]-\mu_i)^2 \qquad , i=1,2 \\ thr_i &= \mu - \sqrt{\sigma_i^2} \end{aligned} \tag{4.20}$$

The thresholded signals, $f_i^{thr}[n]$, are obtained with ease, following Equation 4.21. Note that this technique is intended to find areas of the signal that are significantly below, as measured by the threshold value, the initial one second period, thus identifying ERD.

$$f_i^{thr}[n] = \begin{cases} 1 & \text{if } f_i^{new} \leq thr_i \\ 0 & \text{if } f_i^{new} > 0 \end{cases}, \; i=1,2 \tag{4.21}$$

The various $_i^{thr}[n]$ are combined by computing their mean, and a last threshold is applied, with a task being identified if this mean is higher that 0.5. Following this, a simple post-processing step is performed to ensure that no events are detected outside the time period where subjects were asked to perform the motor task.

The final outcome of this subsection is that it is now possible to select all the feature vectors, i.e., columns of the feature matrices, which will be used as labeled data for training in the classification step.

4.4.2 Support Vector Machine

The main goal of the Support Vector Machine (SVM) classifier is to find the best hyperplane which can discriminate between two classes with the maximum possible margin [28]. Intuitively, in order to separate a scatter of 2D points into two classes, one would define a line dividing the set of points such that all the points belonging to one class are, for instance, on the left side of the line, and all the points belonging the other class are on the right side of the line. To avoid classification errors for unseen data, one would even draw the line as far away as possible from any near points. This is precisely what a SVM classifier does. The line separating the two classes, or more generally, for data in $\mathbb{R}^n$, the hyperplane, is given by:

$$\mathbf{w} \cdot \mathbf{x} - b = 0 \tag{4.22}$$

where **x** are the data points and **w** is the vector orthogonal to the hyperplane. The parameter b determines the offset of the hyperplane from the origin along the normal vector **w**. See Figure 4.7 for a graphical representation.

If the data points are linearly separable, there exist two hyperplanes parallel to the margin such that they are as far away as possible and there are no points between them:

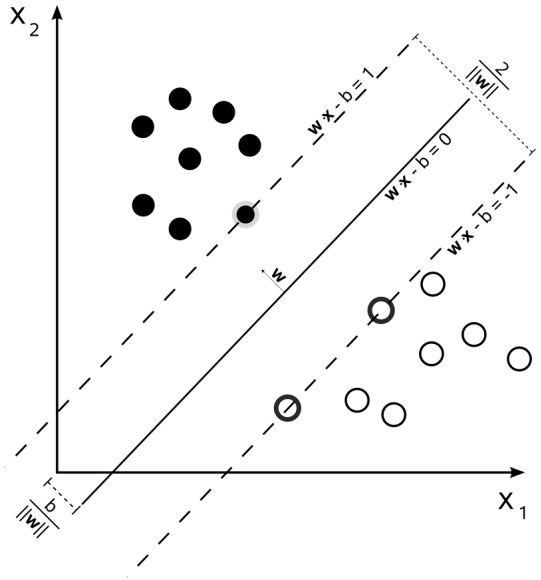


Figure 4.7: Maximum-margin hyperplane for a 2D linear SVM classifier.

$$\mathbf{w} \cdot \mathbf{x} - b = 1 \tag{4.23}$$

$$\mathbf{w} \cdot \mathbf{x} - b = -1 \tag{4.24}$$

The distance between the hyperplanes is $\frac{2}{\|\mathbf{w}\|}$, thus, in order to maximize this distance, it is necessary to minimize $\|\mathbf{w}\|$. Therefore, the problem to optimize is:

$$\underset{\mathbf{w},b}{\operatorname{argmin}} \|\mathbf{w}\| \quad \text{subject to} \tag{4.25}$$

$$(\mathbf{w} \cdot \mathbf{x}_i) d_i - b \geq 1 \tag{4.26}$$

where $d_i = \{-1, 1\}$ is the class of the data point $\mathbf{x}_i$. Note that the vector $\mathbf{w}$, as it belongs to the same subspace of the data, it may be written as a linear combination of the data vectors and, equivalently, as a linear combination of a subset of the data vectors. This subset are the support vectors.

The extension of the linear SVM to nonlinear data implies the definition a transformation that converts the data (feature) space into a space where the problem is linearly separable. As the SVM classifier requires the computation of various internal products, what is usually done is the definition of a kernel function such that the computation of the internal products is easy. This is known as the kernel trick. In this work, a Gaussian kernel was used.

4.4.2.A Implementation

As the SVM classifier used is defined for a two-class problem, a hierarchical approach was taken to classify the entirety of the seven tasks. The hierarchical classifier, as depicted in Figure 4.8, uses a sequence of classifiers, with each distinguishing between two classes. This approach was used for both the power features and the PLF features. Four additional hierarchical classifiers were employed

for the PLF features, as shown in Figure 4.9. These classifiers, while derived from the first, use only five classes. The purpose of this is to assess which of the tasks associated with inferior limbs, moving the feet or the legs, provide a better combination to control a BCI with four channels (remember that task CC is the no-movement task). This type of system could be used, for instance, in a 2D cursor BCI, in which the users would use the four mental tasks to command the cursor (e.g. use the right hand motor task to make the cursor go to the right). The tasks pertaining to the hand movements were always kept because they are easier to detect.

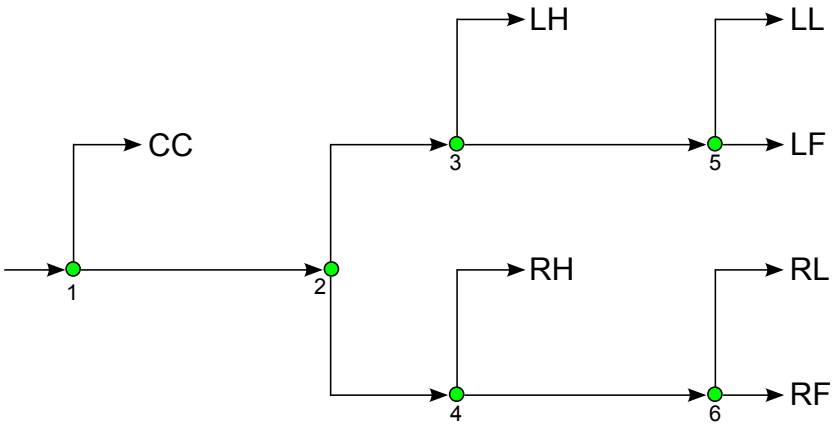
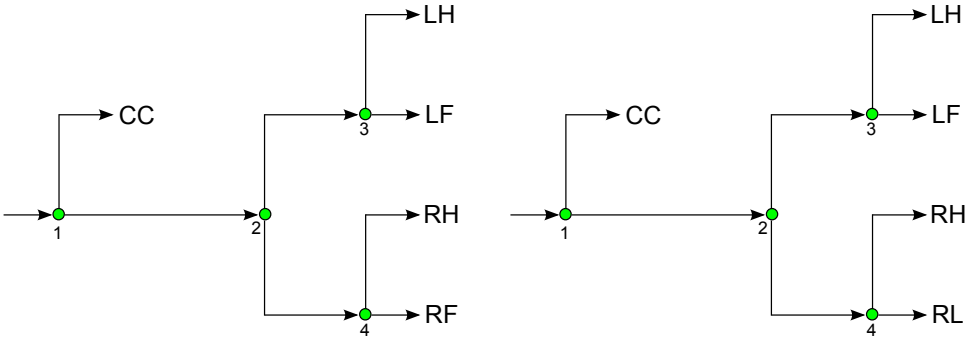


Figure 4.8: Hierarchical SVM classifier for all classes; green dots represent the SVMs.

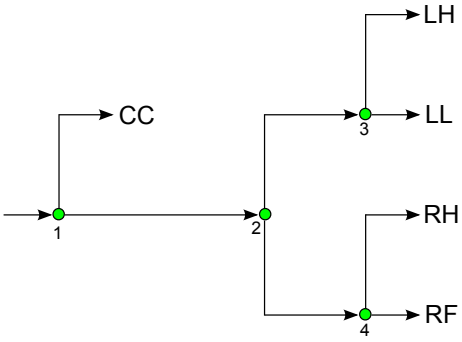
To avoid confusion in the remainder of the dissertation, the classifier in Figure 4.8 will be denoted with upper case (i.e. Hierarchical Classifier), while the alternative classifiers in Figure 4.9 will be written with lower case and followed by the classifier label (e.g. hierarchical classifier H1).

In order to evaluate the performance of the classifiers, the Leave-One-Out Cross Validation method was used. With this method, the training is performed using all but one of the training examples, and the classifier is tested using the excluded sample. This is repeated until all the training examples have been used once for testing. Although the LOOCV method is rather computationally heavy, it allows to train the classifier with the maximum amount of data, and to test it over the entire dataset.

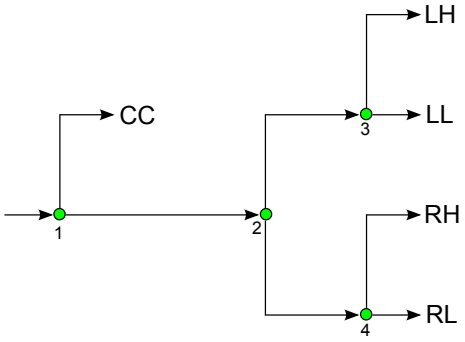


(a) Classifier H1.

(b) Classifier H2.



(c) Classifier H3.



(d) Classifier H4.

Figure 4.9: Alternative hierarchical SVM classifiers ; green dots represent the SVMs.

5

Experimental Results

Contents

In this chapter, a description of the obtained experimental results is presented. This is broken into three sections, with the first section presenting a general overview of the acquired signals. The second section deals with the band power features, and the third covers the PLF features. The last two sections include a quick example of the time-course of each type of features, demonstrating their general properties. This is followed by the classification results obtained for the Hierarchical Classifier. The PLF section also includes the results of the alternative hierarchical classifiers and a simulation of a real-time BCI system using the Hierarchical Classifier.

5.1 The Acquired Signal

To illustrate the type of signals that are dealt with in this thesis, one example is presented in Figure 5.1. In this figure, one can see the EEG signals obtained for one trial of imagined right hand movement, after the preprocessing step, i.e., already bandpass filtered and with the Laplacian Filter applied. It is noticeable that the signals exhibit a sinusoid-like behavior, but the expected ERD is not visible to the naked eye. Also, there is an artifact between seconds 6 and 7, probably caused by eye movements or blinking, as its amplitude decreases from top to bottom (the first 7 channels, on top, are anterior to the remaining channels, thus closer to the eyes).

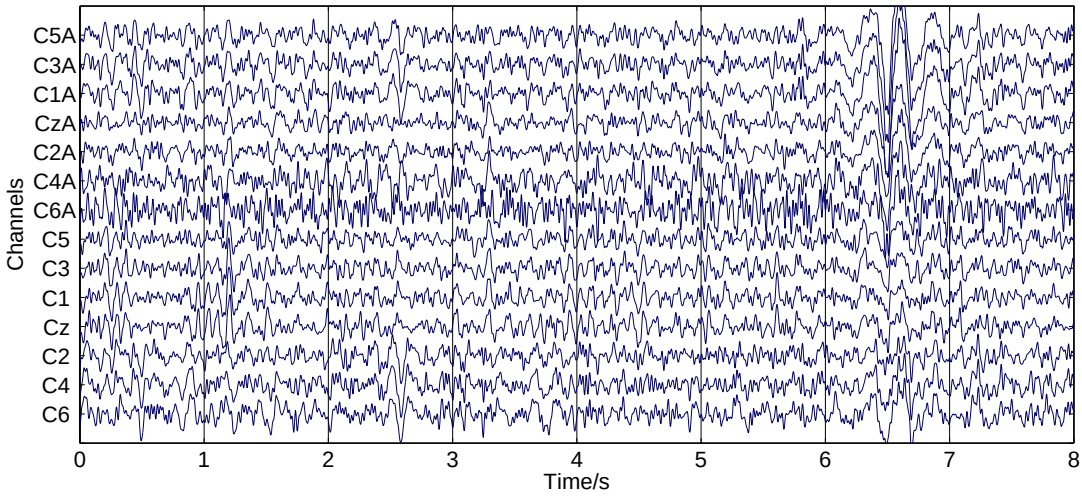
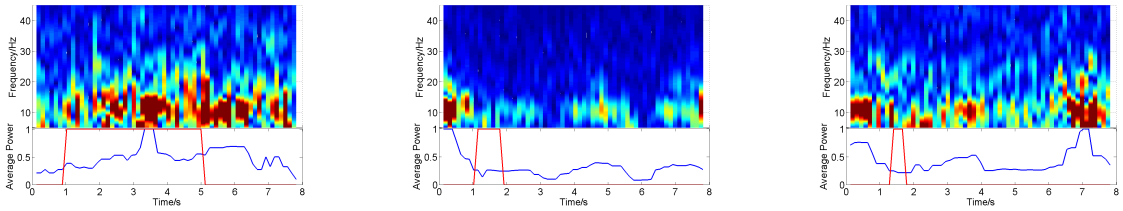


Figure 5.1: Example of one EEG trial corresponding to the imagination of right hand movement, subject 4.

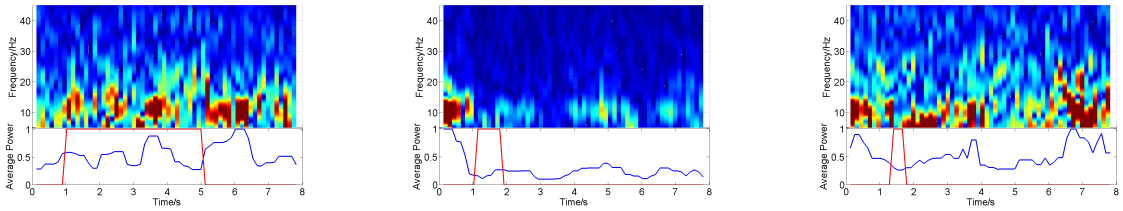
5.2 Band Power Features

5.2.1 The Features

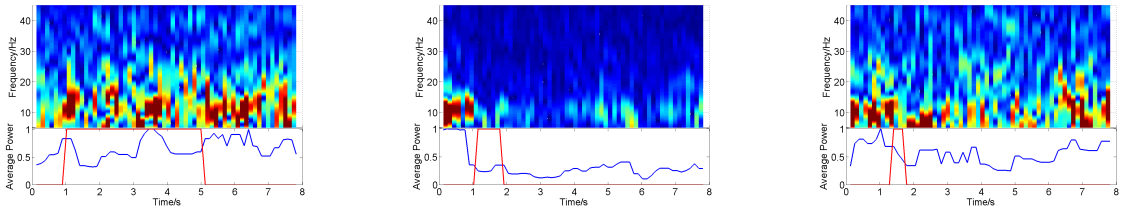
With the goal of understanding how the band power features behave when applied to real data, an illustrative example is presented in Figure 5.2. This figure depicts the spectrograms and average power for the five more important channels over the motor cortex, while the no movement, right hand and left hand tasks were being imagined (corresponding to the three columns of the figure). It is possible to observe that, first of all, there is a defined frequency band around $10\,Hz$ that seems to be the most active across all cases. Second, this frequency band exhibits a decrease in power at the start of the motor task (in the two right columns, after the first second), increasing again at the end of the trial. Third, there is no apparent decrease in power for the no movement task (column to the left), relative to the first second of the trial, although the last two channels do not follow this trend. And, finally, the average power, computed between $8\,Hz$ and $15\,Hz$, sometimes does not show a good agreement with the information present in the spectrogram, making the automatically selected time instants used to train the classifier (in red, on the bottom part of the individual figures) not the best ones. This neatly summarizes what to expect when dealing with EEG acquisitions: either the desired information is clearly visible, or, in other times, it is hidden, and sometimes it is just not present.



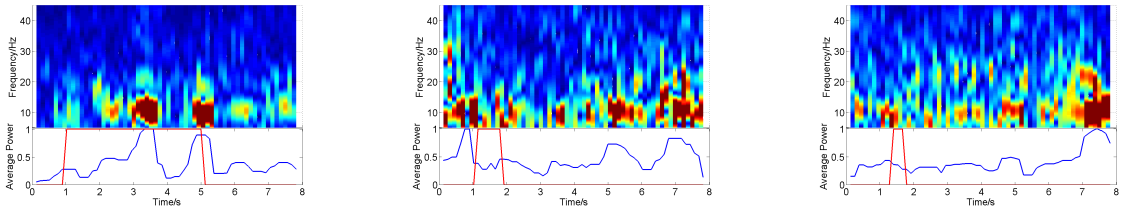
(a) Channel C3.



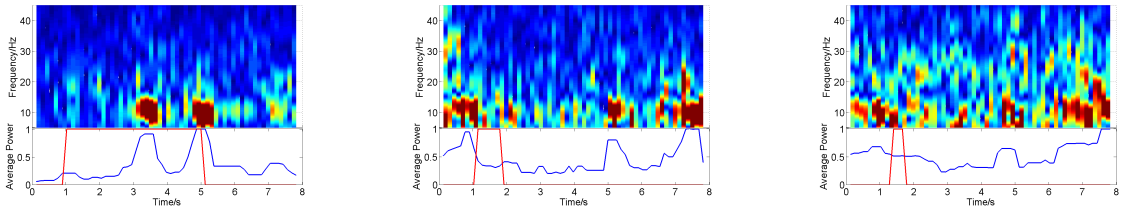
(b) Channel C1.



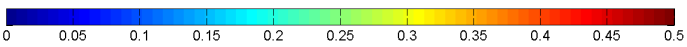
(c) Channel Cz.



(d) Channel C2.



(e) Channel C4.



(f) Color code.

Figure 5.2: Example of the band power features obtained for three imagined tasks performed by subject 4; the columns, from left to right, show Classes 1, 6 and 7 (no movement, right hand, left hand, respectively); in each sub-figure, the top part presents the spectrogram between $5\ Hz$ and $45\ Hz$ and the bottom part plots the average power in blue and the selected time instants to train the classifier in red; the last line shows the color code for the spectrograms, in relative units to the maximum power for each figure.

5.2.2 Hierarchical Classifier

The Hierarchical Classifier, using band power features, was trained, for each subject, with the actual and imagined trials. The global accuracy of the classifier, using the LOOCV method, is presented in Table 5.1. Note that, in this table and all the following, the parameter Rlim. (Random Limit) is used to estimate the expected classification accuracy, if all the individual SVM classifiers in the hierarchy were random classifiers, with a $1/2$ probability of correct classification. This is done using the actual histogram of the classes, i.e., the distribution of the number of examples in the test data, and using Equation 5.1, where $N(i)$ is the number of examples belonging to Class i, and $p(i)$ is the probability of on example of that class being correctly classified, which is defined in Equation 5.2.

$$\text{Rlim} = \frac{1}{\sum_{i=1}^{7} N(i)} \sum_{i=1}^{7} p(i)\, N(i) \tag{5.1}$$

$$p(i) = \begin{cases} \left(\frac{1}{2}\right) & i = 1 \\ \left(\frac{1}{2}\right)^4 & i = 2, 3, 4, 5 \\ \left(\frac{1}{2}\right)^3 & i = 6, 7 \end{cases} \tag{5.2}$$

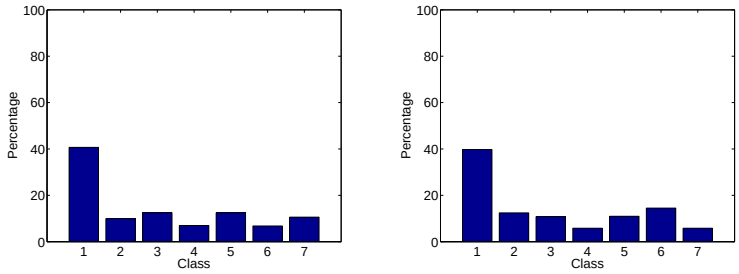
The exponents in Equation 5.2 arise from the number of classifiers that have to be used in order to classify a given class. For instance, to have a test example classified as belonging to Class 7 it is necessary to use three SVM classifiers. Thus, with a $1/2$ probability of correct classification for each SVM, the probability of correctly classifying that class is $\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$. Also, note that, if the class distribution was uniform, the expected accuracy would be $1/7 \approx 14\%$ because, coincidently, the sum of the $p(i)$ over all i is 1. In reality, as will be shown in the next paragraphs, there are more testing examples belonging to Class 1 than to the other classes, resulting in a higher Rlim. value.

Table 5.1: Classification results of the Hierarchical Classifier with band power features, for both the actual and imagined movement tasks; Acc. – Accuracy, Rlim. – Random Limit.

Type		S1	S2	S3	S4	S5	S6	Average
Actual	Acc. (%)	64.24	64.62	75.05	76.81	67.08	64.20	68.67
	Rlim. (%)	24.08	25.13	25.85	24.91	25.15	24.78	24.98
Imagined	Acc. (%)	67.95	65.83	75.92	77.29	74.94	69.25	71.86
	Rlim. (%)	24.86	26.62	27.66	23.50	27.31	24.57	25.75

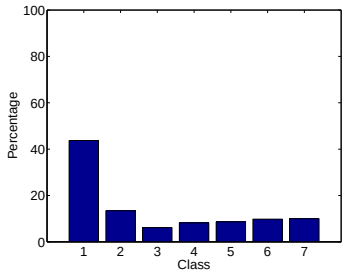
In the remainder of this section, in order to maintain the number of pages not too high, only the results pertaining to two subjects (subjects 2 and 4) are presented.

The class distributions of the testing data for both subjects are shown in Figure 5.3. In this figure, it is possible to observe that, for all the cases, about 40% of the test data belongs to Class 1, i.e., the no movement class, with the remainder of the classes being more or less evenly distributed. This is a direct consequence of how the testing data is generated, as the examples from the other classes are obtained only from the segments that are considered to contain the motor task, according to the threshold defined in the previous chapter.

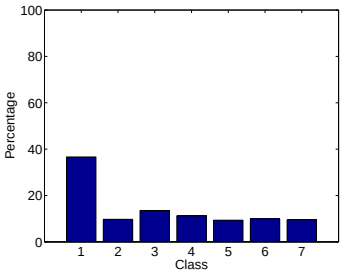


(a) Subject 2, actual motor tasks.

(b) Subject 4, actual motor tasks.



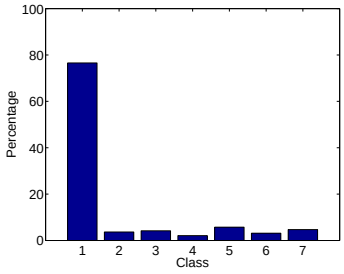
(c) Subject 2, imagined motor tasks.



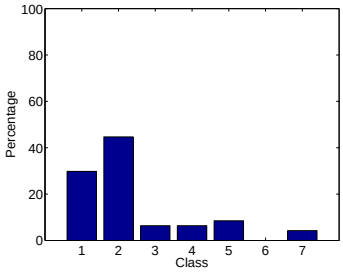
(d) Subject 4, imagined motor tasks.

Figure 5.3: Class distribution of the training set for the Hierarchical Classifier with band power features.

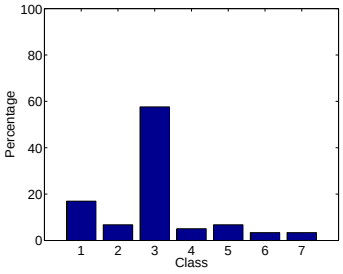
Figure 5.4 presents the classification results for each class, i.e., how examples from a given motor tasks are classified, for subject 2 performing actual motor tasks. It is visible that, for elements of a given class, they are more often correctly classified than otherwise. Despite this, Classes 2, 3, 6, and 7 exhibit significant classification errors, in particular mistaking these classes for Class 1. Similar results are obtained for the imagined tasks, presented in Figure 5.5, with the addition that here Classes 2 and 4 have classification accuracies below 50%.



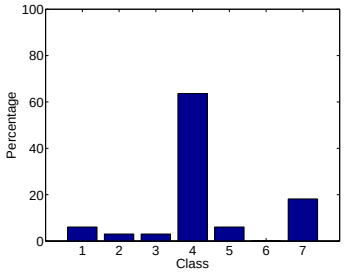
(a) Class 1: 73.44% acc.



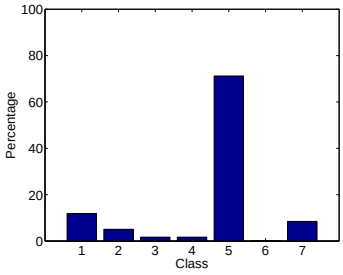
(b) Class 2: 53.06% acc.



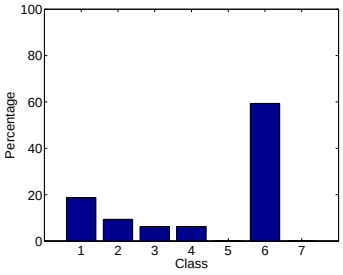
(c) Class 3: 51.02% acc.



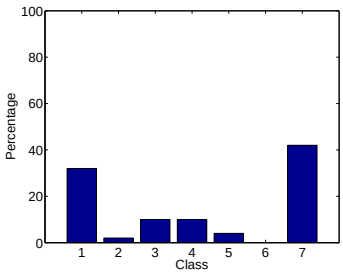
(d) Class 4: 68.29% acc.



(e) Class 5: 58.18% acc.

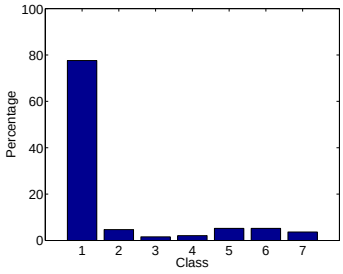


(f) Class 6: 56.52% acc.

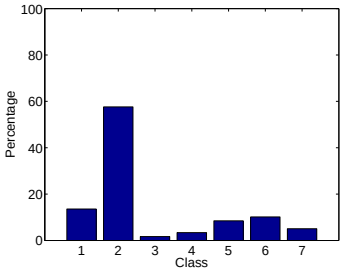


(g) Class 7: 54.55% acc.

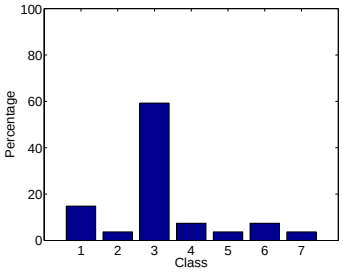
Figure 5.4: Accuracy results for each of the classes in the Hierarchical Classifier with band power features, for subject 2, actual motor tasks; acc – accuracy.



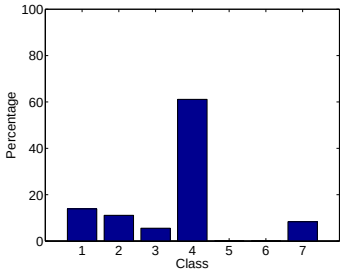
(a) Class 1: 78.65% acc.



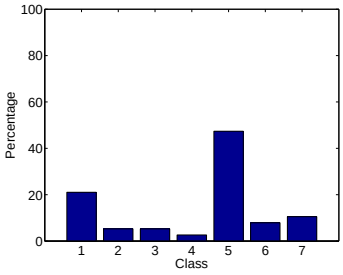
(b) Class 2: 42.86% acc.



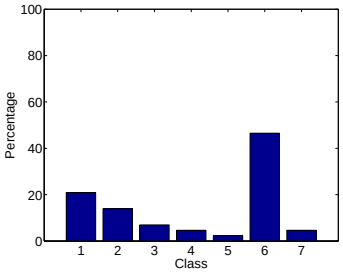
(c) Class 3: 63.64% acc.



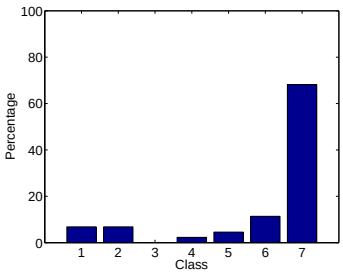
(d) Class 4: 47.06% acc.



(e) Class 5: 71.93% acc.



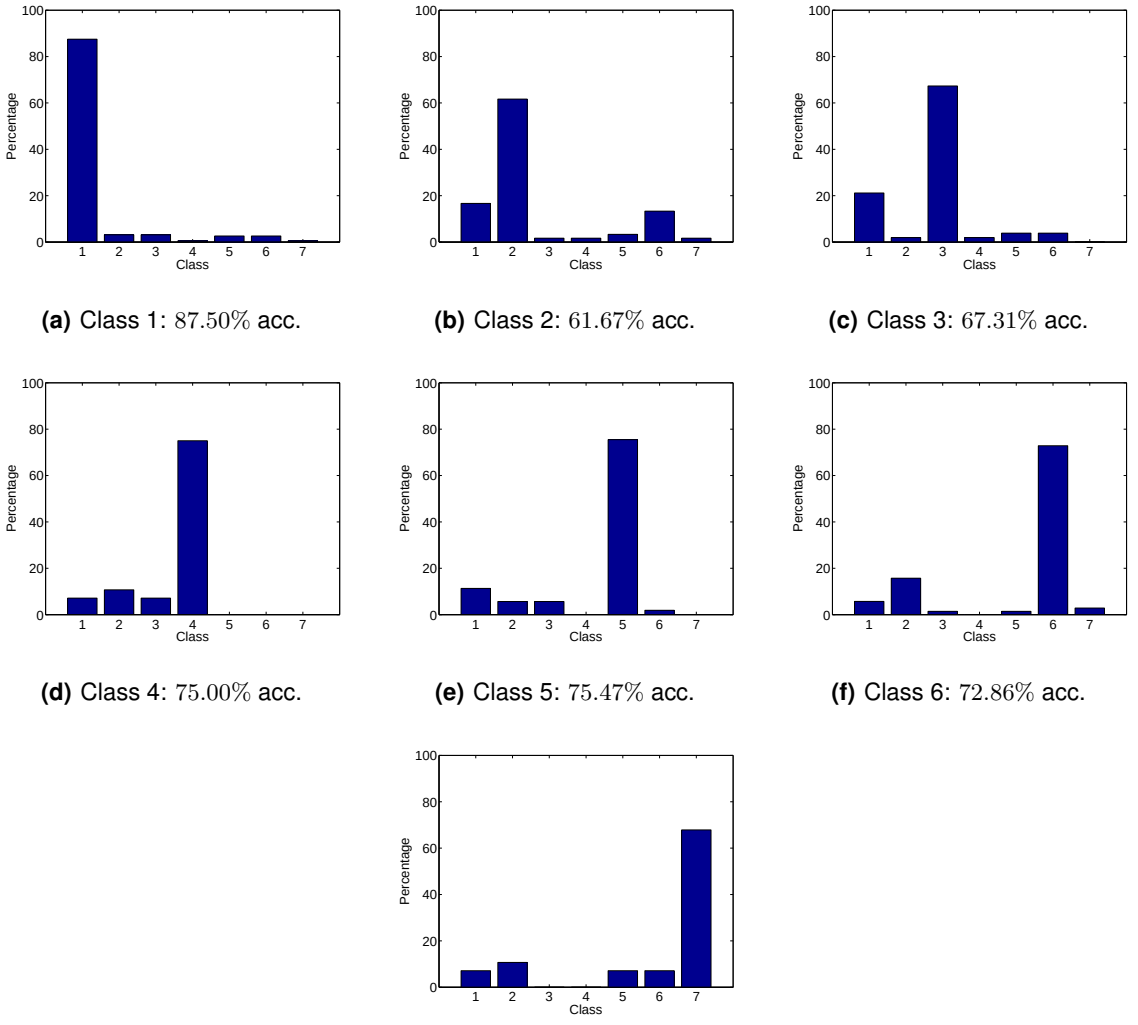
(f) Class 6: 56.52% acc.



(g) Class 7: 76.36% acc.

Figure 5.5: Accuracy results for each of the classes in the Hierarchical Classifier with band power features, for subject 2, imagined motor tasks; acc – accuracy.

Figures 5.6 and 5.7 present again the classification results for each class, but now pertaining to subject 4. It is evident a general increase in accuracy, when compared to subject 2, although the confusion with Class 1 persists for the Classes 2 and 3, for the actual motor tasks. This is not as noticeable in the imagined tasks, where there is a slight increase in classification accuracy. Actually, the increase in accuracy from the actual to the imagined tasks is also true for all the subjects, as is discernible in Table 5.1. This comes a bit as a surprise, given that the imagined motor tasks are expected to produce less noticeable effects in the EEG than the actual motor tasks.



(a) Class 1: 87.50% acc.

(b) Class 2: 61.67% acc.

(c) Class 3: 67.31% acc.

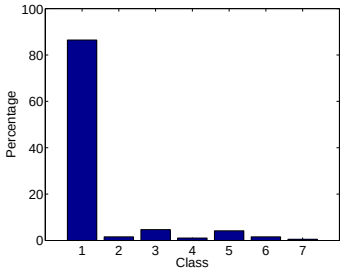
(d) Class 4: 75.00% acc.

(e) Class 5: 75.47% acc.

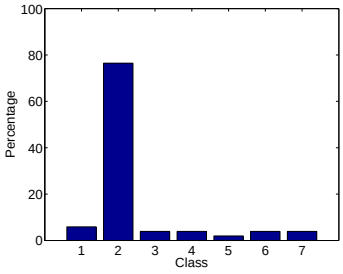
(f) Class 6: 72.86% acc.

(g) Class 7: 67.86% acc.

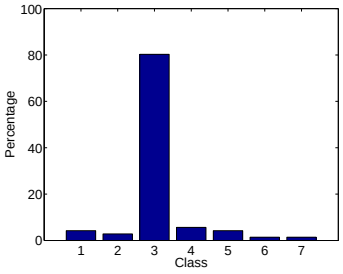
Figure 5.6: Accuracy results for each of the classes in the Hierarchical Classifier with band power features, for subject 4, actual motor tasks; acc – accuracy.



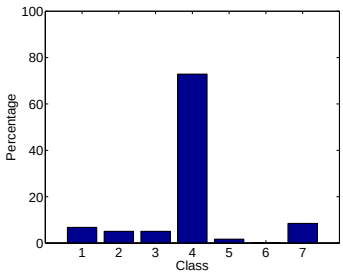
(a) Class 1: 86.46% acc.



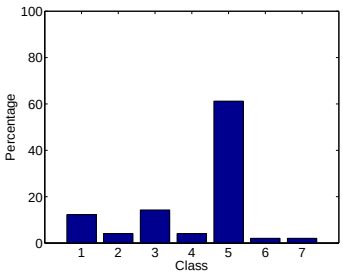
(b) Class 2: 76.47% acc.



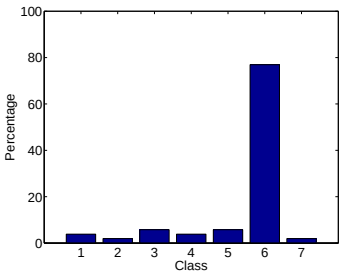
(c) Class 3: 80.28% acc.



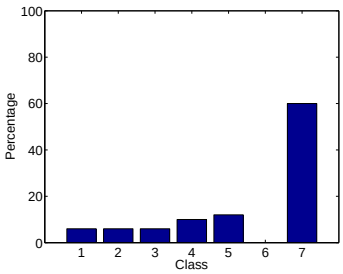
(d) Class 4: 72.88% acc.



(e) Class 5: 61.22% acc.



(f) Class 6: 76.92% acc.



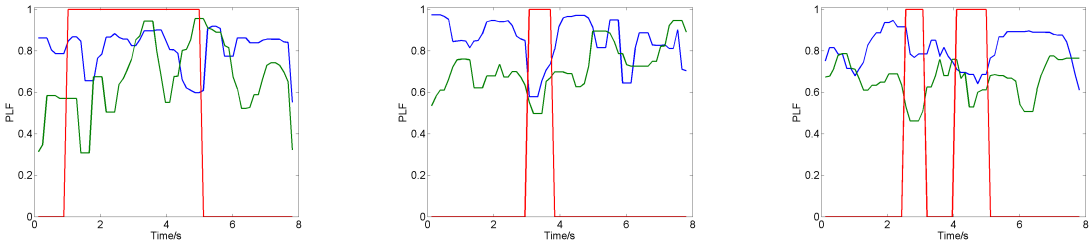
(g) Class 7: 60.00% acc.

Figure 5.7: Accuracy results for each of the classes in the Hierarchical Classifier with band power features, for subject 4, imagined motor tasks; acc – accuracy.

5.3 Phase-Locking Factor Features

5.3.1 The Featues

Similarly to what was done with the band power features, an illustrative example of the general properties of the PLF features is presented in Figure 5.8. It is apparent in this figure, where the PLF between channels C3-C1 is shown in blue and the PLF between channels C4-C6 is in green, that these features are easier to interpret than the band power features. First, in Classes 6 and 7, it is visible that the features exhibit a somewhat smooth evolution in the first two seconds, followed by a sudden drop. This marks the beginning of the motor task. Second, while in Class 6 the C3-C1 PLF has a sharp decrease and the C4-C6 PLF has a less evident drop, the contrary happens for Class 7, thus providing a straightforward way to distinguish the two classes. Finally, note that, in Class 7, the second segment of selected time instants (in red) does not follow this rule exactly. In all honesty, if the same thresholding technique was used for Class 1, there would be segments in that class identified as being motor tasks, despite that no motor task was performed. Again, this proves that the thresholding method is not the best, although it provides better results for the PLF features than for the band power features.



(a) Class 1. **(b)** Class 6. **(c)** Class 7.

Figure 5.8: Example of the PLF features obtained for three imagined tasks performed by subject 4; in each sub-figure, the two time courses of the PLF features used to define the selecting threshold are shown (in blue the PLF between channels C3-C1 and in green the PLF between channels C4-C6); the selected time instants used for the classifier are plotted in red.

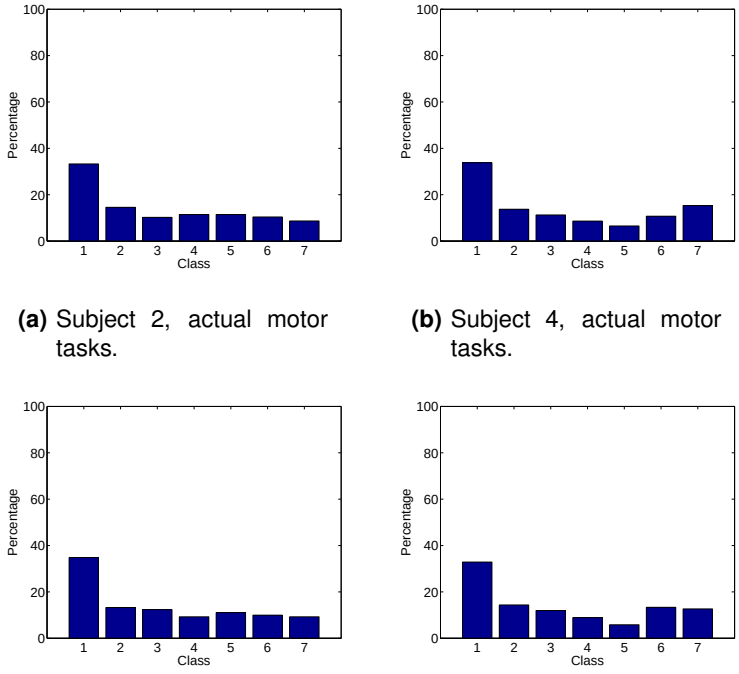
5.3.2 Hierarchical Classifier

The Hierarchical Classifier, now using the PLF features, was trained, for each subject, with the actual and imagined trials. The global accuracy of the classifier, using the LOOCV method, is presented in Table 5.2.

Table 5.2: Classification results of the Hierarchical Classifier with PLF features, for both the actual and imagined movement tasks; Acc. – Accuracy, Rlim. – Random Limit.

Type		S1	S2	S3	S4	S5	S6	Average
Actual	Acc. (%)	86.99	87.18	89.67	84.86	86.45	84.31	86.58
	Rlim. (%)	21.81	22.00	21.47	22.67	22.70	21.71	22.06
Imagined	Acc. (%)	87.54	86.03	87.38	86.30	85.74	85.06	86.34
	Rlim. (%)	21.33	22.70	20.81	22.26	22.58	21.26	21.82

The class distributions of the testing data for both subjects are presented in Figure 5.9. In this figure, it is possible to observe that, for all the cases, about 35% of the test data belongs to Class 1, with the remainder of the classes being more or less evenly distributed. This accounts for a better distribution of the data, when compared with the results obtained for the band power features. Therefore, the Rlim. values are lower than for the band power features, making any change in accuracy more significant.



(a) Subject 2, actual motor tasks.

(b) Subject 4, actual motor tasks.

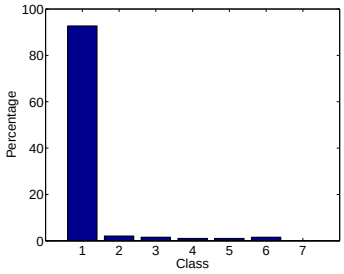
(c) Subject 2, imagined motor tasks.

(d) Subject 4, imagined motor tasks.

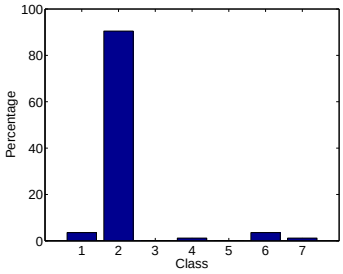
Figure 5.9: Class distribution of the training set for the Hierarchical Classifier with PLF features.

Figure 5.10 presents the classification results for each class, for subject 2 performing imagined motor tasks. It is noticeable that, for elements of a given class, they are clearly more often correctly classified than otherwise. Also, the confusion related to Class 1 is no longer as evident as in the case of the band features. It is also interesting to note that classes pertaining to movements on the right side of the body are less likely to be mistaken for classes pertaining to movements on the left side of the body, and vice-versa. For instance, Class 2 (right foot) is less mistaken for Classes 3, 5, and 7 (left foot, leg and hand, respectively) than it is for Classes 4 and 6 (right leg and hand, respectively).

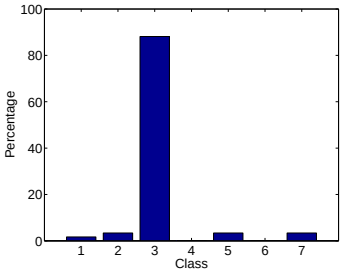
Similar results are obtained for the imagined tasks, presented in Figure 5.11, although there is a slight decrease in accuracy.



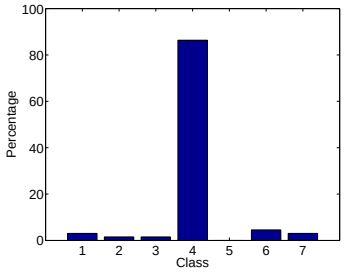
(a) Class 1: 92.71% acc.



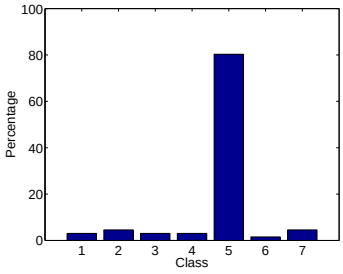
(b) Class 2: 90.48% acc.



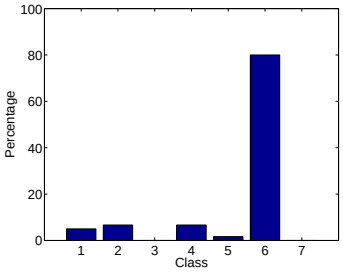
(c) Class 3: 88.14% acc.



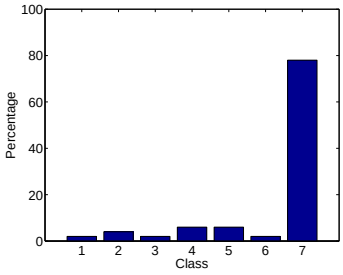
(d) Class 4: 86.36% acc.



(e) Class 5: 80.30% acc.

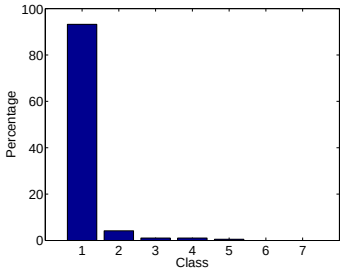


(f) Class 6: 80.00% acc.

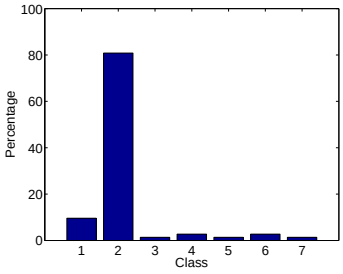


(g) Class 7: 78.00% acc.

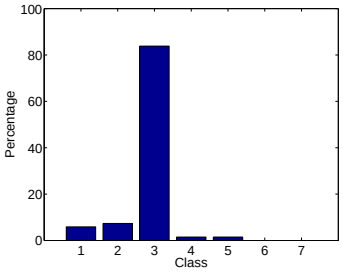
Figure 5.10: Accuracy results for each of the classes in the Hierarchical Classifier with PLF features, for subject 2, actual motor tasks; acc – accuracy.



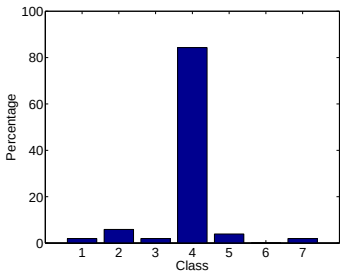
(a) Class 1: 93.23% acc.



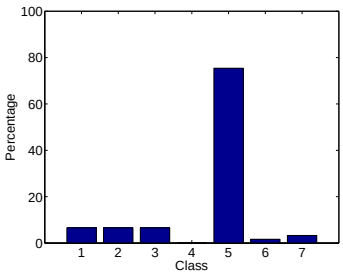
(b) Class 2: 80.82% acc.



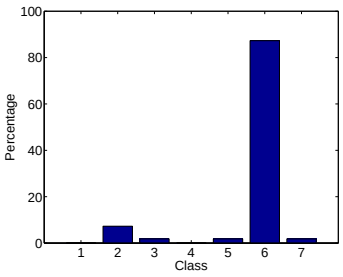
(c) Class 3: 83.82% acc.



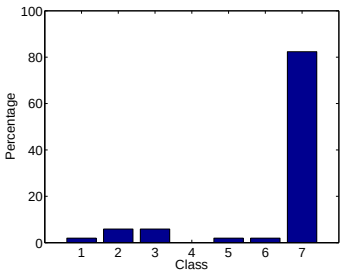
(d) Class 4: 84.31% acc.



(e) Class 5: 75.41% acc.



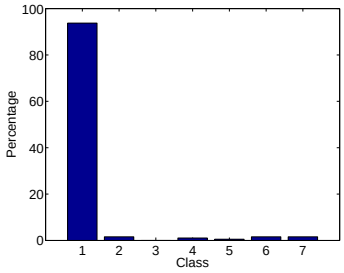
(f) Class 6: 87.27% acc.



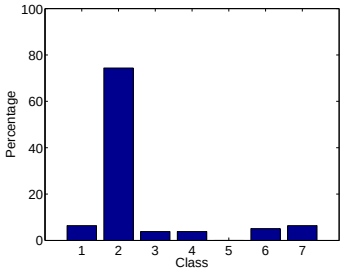
(g) Class 7: 82.35% acc.

Figure 5.11: Accuracy results for each of the classes in the Hierarchical Classifier with PLF features, for subject 2, imagined motor tasks; acc – accuracy.

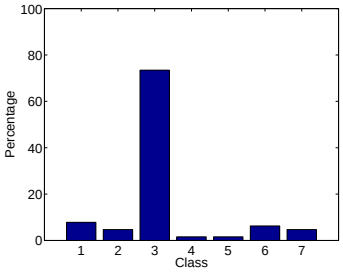
Figures 5.12 and 5.13 present the classification results for each class, but now pertaining to subject 4. While the actual motor tasks show a decrease in accuracy relative to subject 2, the results obtained for the imagined tasks are very similar. In fact, it is evident from Table 5.2 that the variance between the subjects for the PLF features is small (lower than 2%), and the average accuracy is very similar between the actual and the imagined tasks.



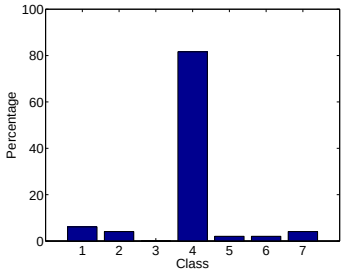
(a) Class 1: 93.75% acc.



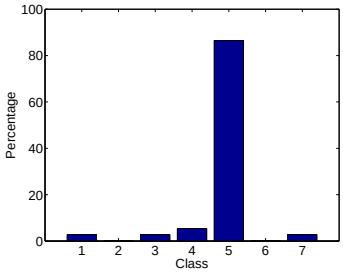
(b) Class 2: 74.36% acc.



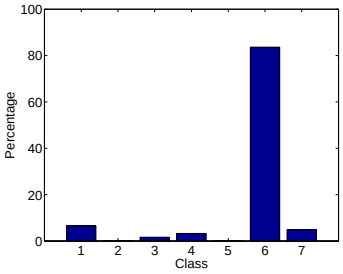
(c) Class 3: 73.44% acc.



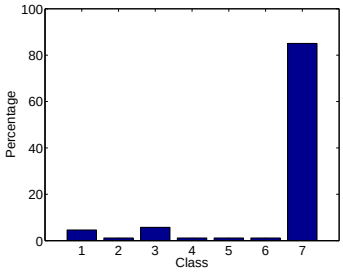
(d) Class 4: 81.63% acc.



(e) Class 5: 86.49% acc.

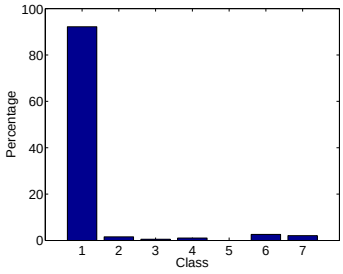


(f) Class 6: 83.61% acc.

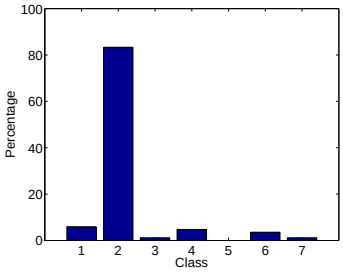


(g) Class 7: 85.06% acc.

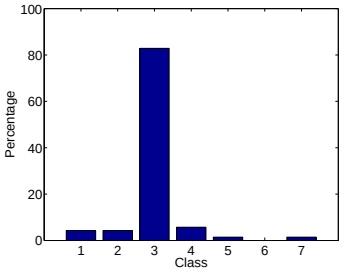
Figure 5.12: Accuracy results for each of the classes in the Hierarchical Classifier with PLF features, for subject 4, actual motor tasks; acc – accuracy.



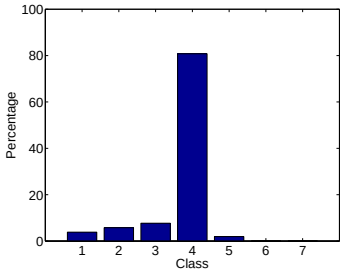
(a) Class 1: 92.19% acc.



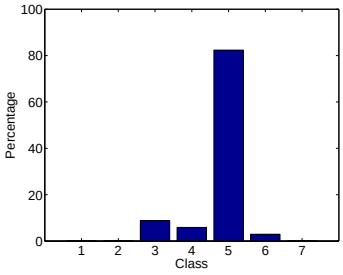
(b) Class 2: 83.33% acc.



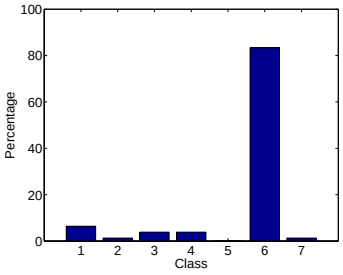
(c) Class 3: 82.86% acc.



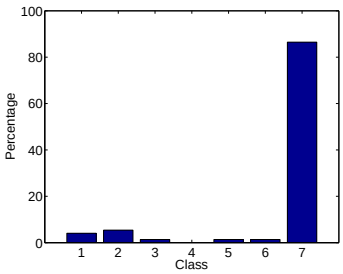
(d) Class 4: 80.77% acc.



(e) Class 5: 82.35% acc.



(f) Class 6: 83.33% acc.



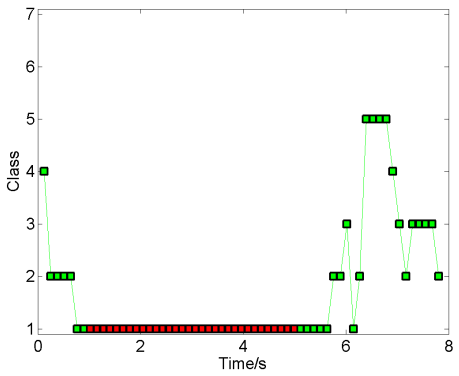
(g) Class 7: 86.49% acc.

Figure 5.13: Accuracy results for each of the classes in the Hierarchical Classifier with PLF features, for subject 4, imagined motor tasks; acc – accuracy.

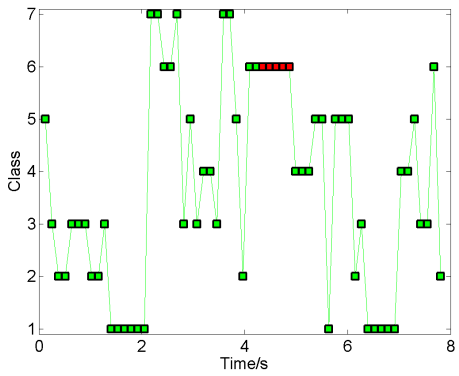
5.3.2.A Simulation of Real-Time Classification

In order to simulate the application of the Hierarchical Classifier in a real-time BCI system (with the PLF features, as they provided better results), the classifier was retrained with all the available data. The resulting classifier was then applied to the entirety of selected trials, as is shown in Figures 5.14 and 5.15.

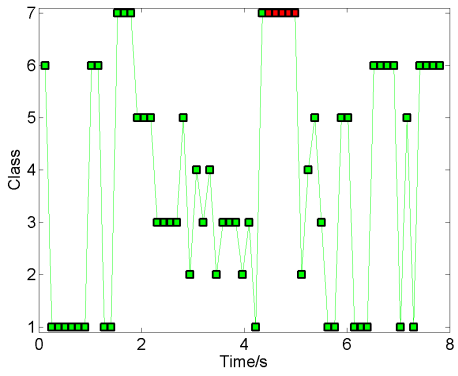
The ideal behavior would be to have a baseline of Class 1 (no movement), with temporary transitions to the correct class. Clearly, this is not the case. Keeping in mind that from the fifth second onwards is the relax period and, therefore, anything can happen, it seems that Class 1 is reasonably well classified in its own trials. As for the other two classes, a lot more jitter is present, in particular for Class 6 in Figure 5.14. On the other hand, the same class but in Figure 5.15 is very close to the ideal trail.



(a) Class 1.

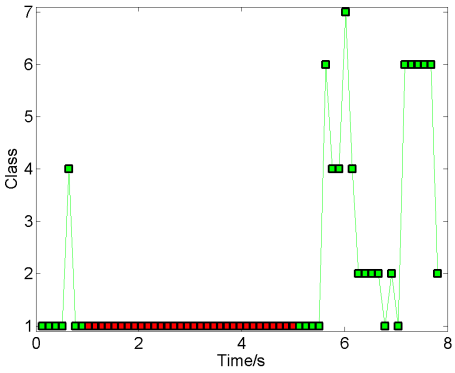


(b) Class 6.

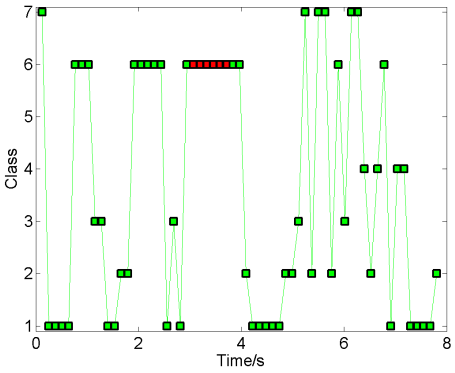


(c) Class 7.

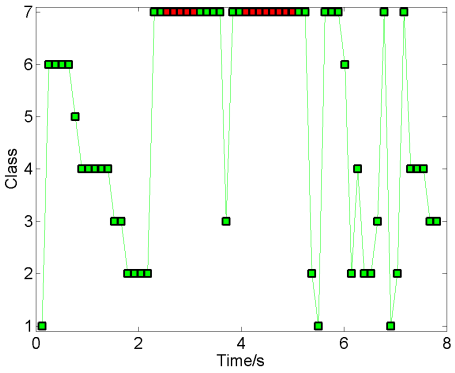
Figure 5.14: Application of the Hierarchical Classifier trained with PLF features to trials from three classes for subject 2 performing imagined motor tasks; the green squares identify the class predicted by the classifier and the red squares identify training data.



(a) Class 1.



(b) Class 6.



(c) Class 7.

Figure 5.15: Application of the Hierarchical Classifier trained with PLF features to trials from three classes for subject 4 performing imagined motor tasks; the green squares identify the class predicted by the classifier and the red squares identify training data.

5.3.3 Alternative Classifiers

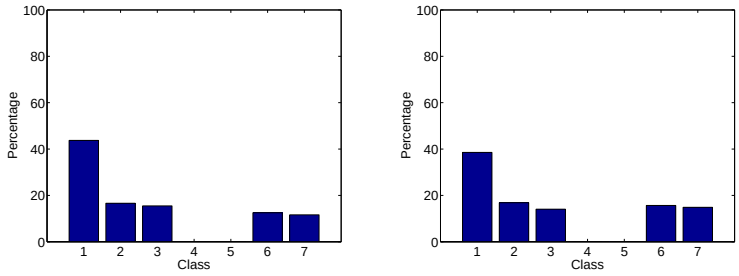
The accuracy results obtained for the alternative classifiers are shown in Table 5.3. It is discernible a general increase in accuracy when compared to the Hierarchical Classifier with PLF features.

Table 5.3: Classification results of the alternative hierarchical classifiers with PLF features, for imagined motor tasks; Acc. – Accuracy, Rlim. – Random Limit.

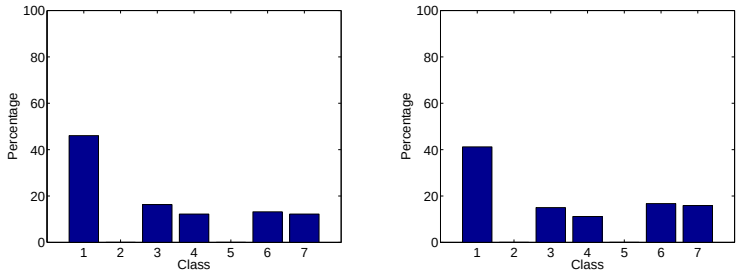
Classifier		S1	S2	S3	S4	S5	S6	Average
H1	Acc. (%)	89.05	89.98	90.74	86.95	88.91	88.48	89.02
	Rlim. (%)	25.50	26.89	25.49	25.03	26.98	26.35	26.04
H2	Acc. (%)	89.09	90.17	90.20	90.77	90.23	87.58	89.67
	Rlim. (%)	25.08	27.98	25.33	26.31	27.56	26.04	26.38
H3	Acc. (%)	87.88	87.33	88.51	89.18	88.09	89.92	88.49
	Rlim. (%)	26.42	27.23	25.10	26.49	26.84	25.67	26.29
H4	Acc. (%)	89.64	89.76	89.60	91.40	87.76	89.44	89.60
	Rlim. (%)	25.95	28.35	24.95	27.99	27.41	25.39	26.67

The class distributions of the testing data for both subjects are presented in Figure5.16, where it is possible to observe that, for all cases, about 40% of the test data belongs to Class 1. This produces higher Rlim. values than before, as can be seen in Table 5.3. Also note that, as only five of the classes are being used in these alternative classifiers, the remaining classes are obviously not used to train or test the classifiers.

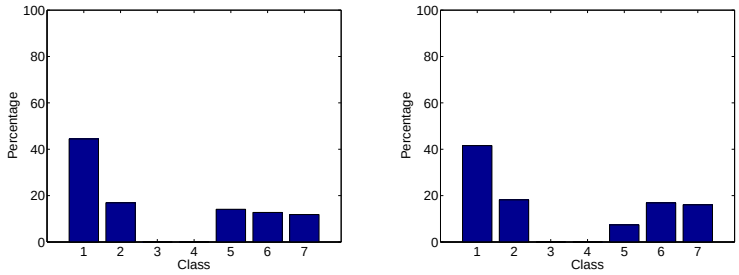
In Figures 5.17 to 5.24 the classification results for the four alternative classifiers are shown for both subjects performing imagined tasks. All the cases produce identical results, with no statistically significant differences, not being possible to ascertain which classifier is the best. Despite this, slightly better results are obtained for classifier H2, which combines movements of both hands with movements of the left foot and the right leg. On the other hand, the worst classifier is the classifier H3 (both hands, left leg and right foot), although for subject 4 the worst classifier is the classifier H1.



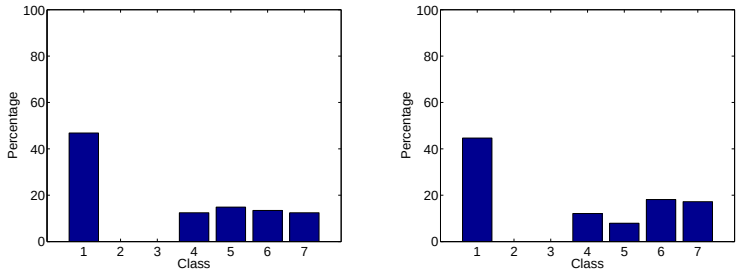
(a) Subject 2, classifier H1. **(b)** Subject 4, classifier H1.



(c) Subject 2, classifier H2. **(d)** Subject 4, classifier H2.

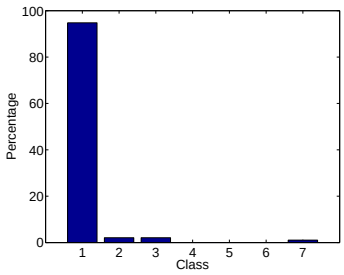


(e) Subject 2, classifier H3. **(f)** Subject 4, classifier H3.

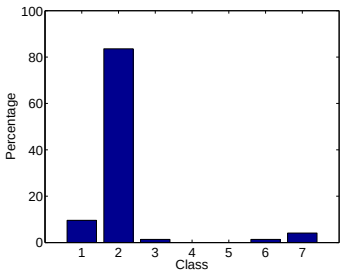


(g) Subject 2, classifier H4. **(h)** Subject 4, classifier H4.

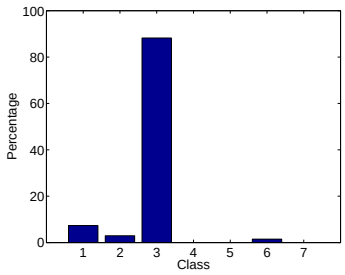
Figure 5.16: Class distribution of the training set for the alternative hierarchical classifiers with PLF features.



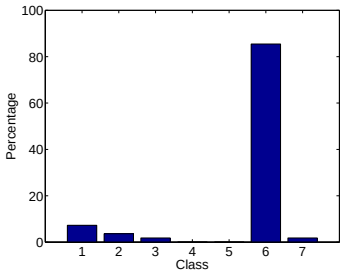
(a) Class 1: 94.79% acc.



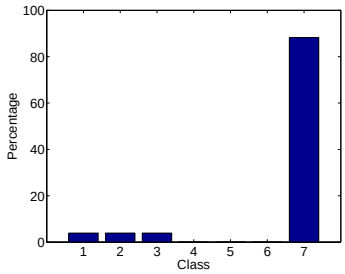
(b) Class 2: 83.56% acc.



(c) Class 3: 88.24% acc.

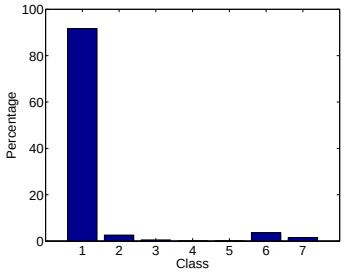


(d) Class 6: 85.45% acc.

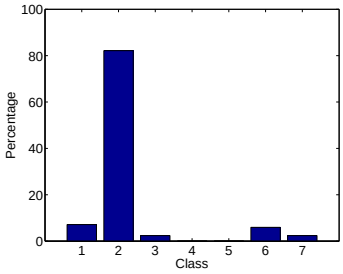


(e) Class 7: 88.24% acc.

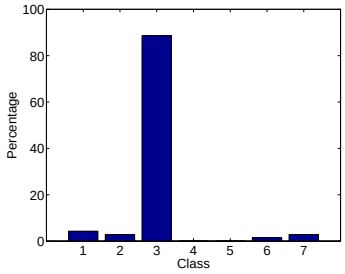
Figure 5.17: Accuracy results for each of the classes in the hierarchical classifier H1 with PLF features, for subject 2, imagined motor tasks; acc – accuracy.



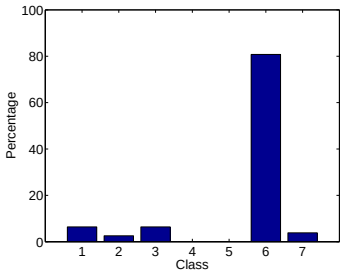
(a) Class 1: 91.67% acc.



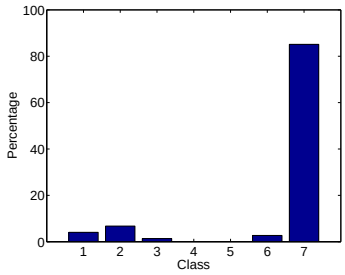
(b) Class 2: 82.14% acc.



(c) Class 3: 88.57% acc.

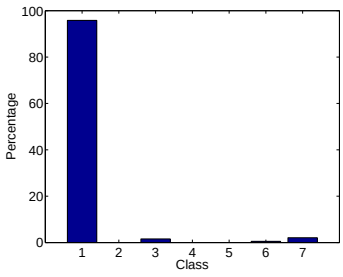


(d) Class 6: 80.77% acc.

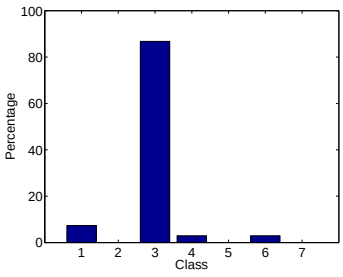


(e) Class 7: 85.14% acc.

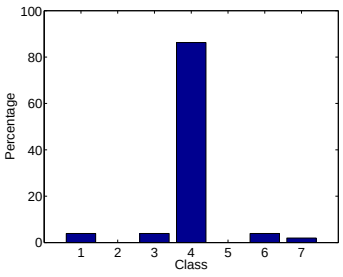
Figure 5.18: Accuracy results for each of the classes in the hierarchical classifier H1 with PLF features, for subject 4, imagined motor tasks; acc – accuracy.



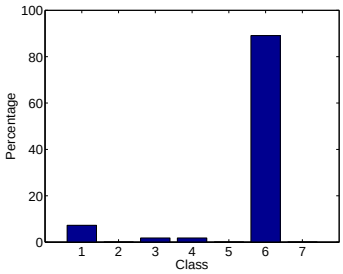
(a) Class 1: 95.83% acc.



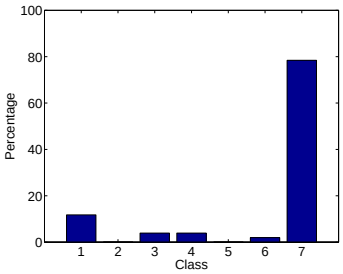
(b) Class 3: 86.76% acc.



(c) Class 4: 86.27% acc.

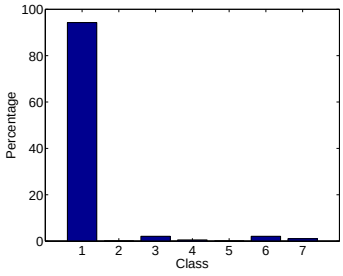


(d) Class 6: 89.09% acc.

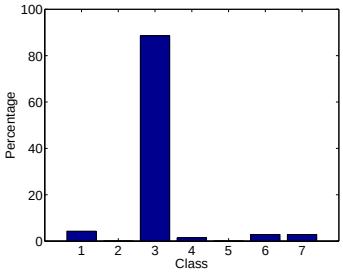


(e) Class 7: 78.43% acc.

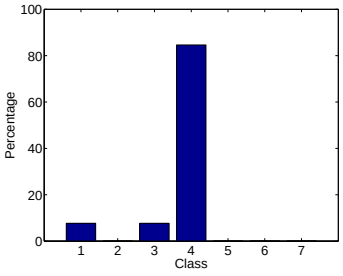
Figure 5.19: Accuracy results for each of the classes in the hierarchical classifier H2 with PLF features, for subject 2, imagined motor tasks; acc – accuracy.



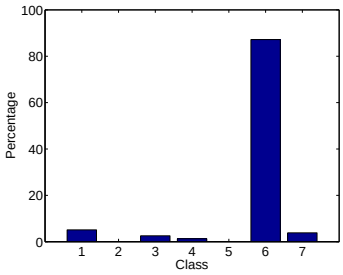
(a) Class 1: 94.27% acc.



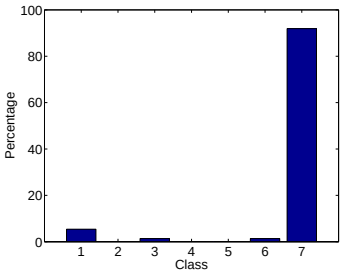
(b) Class 3: 88.57% acc.



(c) Class 4: 84.62% acc.

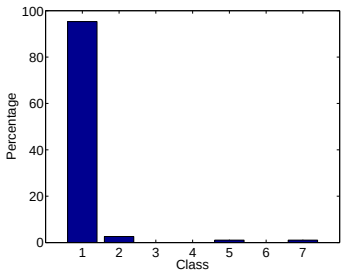


(d) Class 6: 87.18% acc.

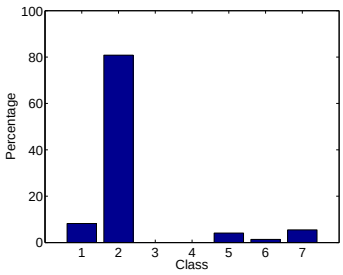


(e) Class 7: 91.89% acc.

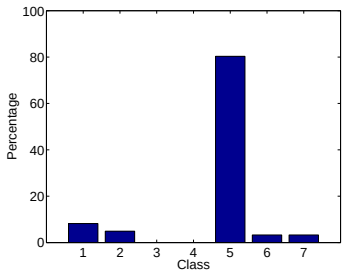
Figure 5.20: Accuracy results for each of the classes in the hierarchical classifier H2 with PLF features, for subject 4, imagined motor tasks; acc – accuracy.



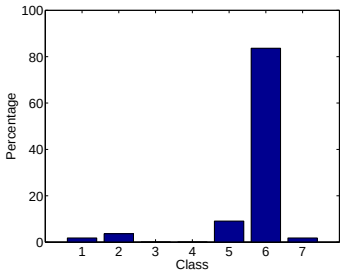
(a) Class 1: 95.31% acc.



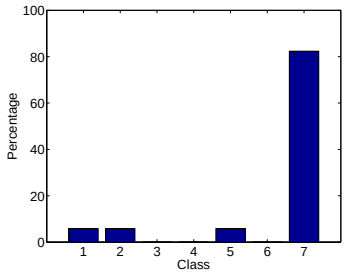
(b) Class 2: 80.82% acc.



(c) Class 5: 80.33% acc.

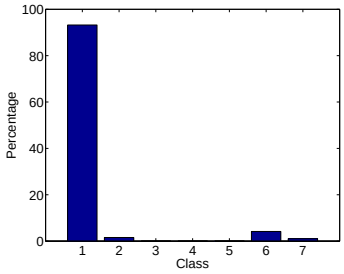


(d) Class 6: 83.64% acc.

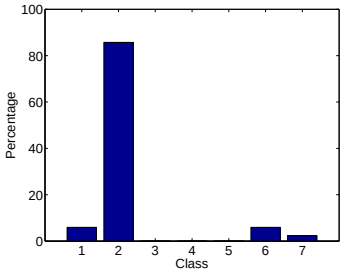


(e) Class 7: 82.35% acc.

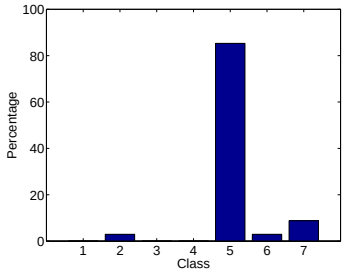
Figure 5.21: Accuracy results for each of the classes in the hierarchical classifier H3 with PLF features, for subject 2, imagined motor tasks; acc – accuracy.



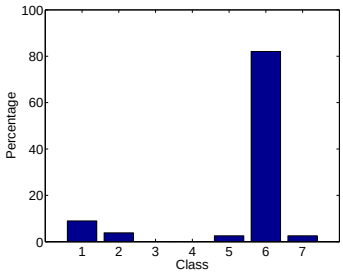
(a) Class 1: 93.23% acc.



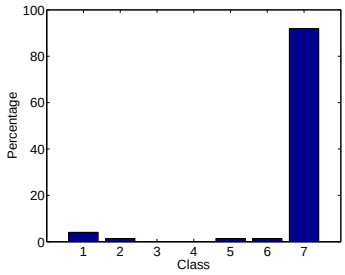
(b) Class 2: 85.71% acc.



(c) Class 5: 85.29% acc.

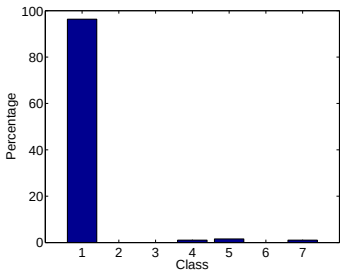


(d) Class 6: 82.05% acc.

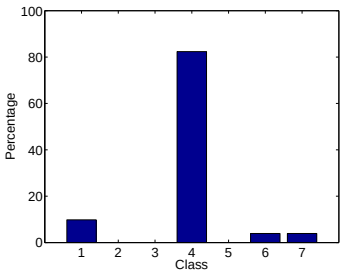


(e) Class 7: 91.89% acc.

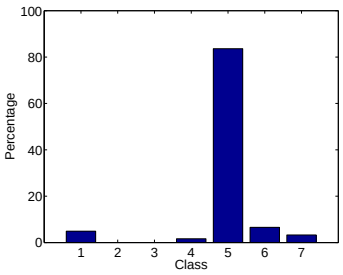
Figure 5.22: Accuracy results for each of the classes in the hierarchical classifier H3 with PLF features, for subject 4, imagined motor tasks; acc – accuracy.



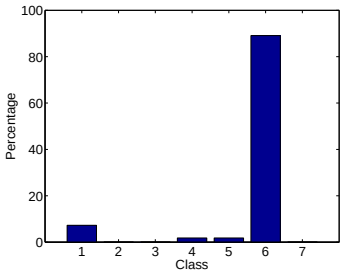
(a) Class 1: 96.35% acc.



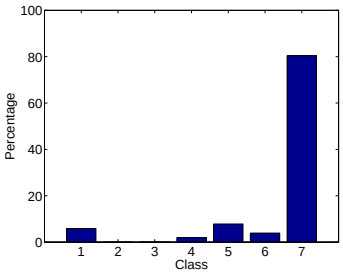
(b) Class 4: 82.35% acc.



(c) Class 5: 83.61% acc.

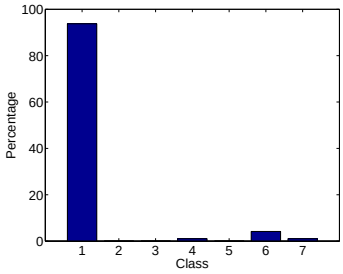


(d) Class 6: 89.09% acc.

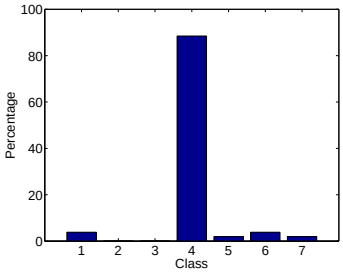


(e) Class 7: 80.39% acc.

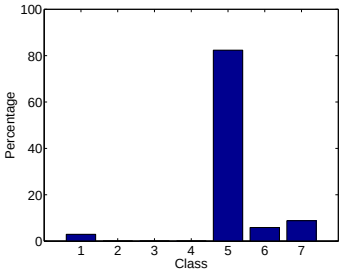
Figure 5.23: Accuracy results for each of the classes in the hierarchical classifier H4 with PLF features, for subject 2, imagined motor tasks; acc – accuracy.



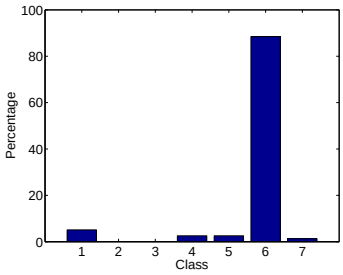
(a) Class 1: 93.75% acc.



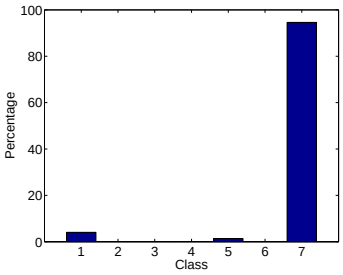
(b) Class 4: 88.46% acc.



(c) Class 5: 82.35% acc.



(d) Class 6: 88.46% acc.



(e) Class 7: 94.59% acc.

Figure 5.24: Accuracy results for each of the classes in the hierarchical classifier H4 with PLF features, for subject 4, imagined motor tasks; acc – accuracy.

6

Discussion

Contents

This chapter summarizes and discusses the experimental results presented in the previous chapter. It also includes proposals for future work and the closing remarks.

6.1 Summary

This thesis sought to develop all the signal processing tools in order to design a Brain-Computer Interface based on the imagination of motor tasks. Having identified the appropriate physiological variables to measure, namely the Event-Related Desynchronization, two different approaches were used: the more traditional method based on band power features and a new method based on the Phase-Locking Factor. These features were then used to train hierarchical Support Vector Machine classifiers and their performance was measured using the Leave-One-Out Cross Validation method.

6.2 Discussion of Results

6.2.1 The EEG Acquisition

The process of acquiring the necessary EEG data was a challenge in itself, in the initial period of the development of this thesis. It encompassed familiarizing with foreign hardware and software, dealing with the correct placement of the electrodes' cap and adjusting the electrodes' impedance individually with conductive gel. Necessarily, the acquisition conditions were different for each subject, resulting in signals with disparate quality. More importantly, the electrodes were probably not placed is the exact same cerebral locations from subject to subject, although the methods proposed here rely on specific electrodes. These problems may only be overcome with practice.

In relation to the experimental procedure itself, it showed some drawbacks that hindered the quality of the results. In particular, the length of the initial period, between the appearance of the fixation cross and the beginning of the motor task, should have been bigger, allowing for the establishment of a better baseline signal to serve as reference for the rest of the trial. Additionally, the motor task period should have been smaller, thus eliminating the need to locate were in this period a motor task is really present. Alternatively, a video recording of the acquisition session could provide this information, although it would only work for the actual motor tasks. Also, more care should have been taken with the description of the motor tasks to be performed by the subjects, maintaining a more uniform mental task across the subjects. This could be done by showing the subjects a video of the desired movements prior to the acquisition. As a side note, it is discussed in [21] that the observation of motor tasks elicits a similar brain response as the actual performance or imagination of the motor tasks.

As a final remark in this matter, it is important to remember that no artifacts were removed from the signals, and all the trails were used, independently from the fact that they may or may not contain artifacts. This obviously reduces the performance of the system.

6.2.2 Comparison Between Band Power Features and PLF Features

The results shown in Chapter 5 allow to conclude that the PLF features are better, with an average accuracy of 86.6% and 86.3%, for the actual and imagined tasks, respectively, against an average accuracy of 68.7% and 71.9% for the band power features. This difference is due to the more robust theoretical formulation of the PLF features, which appear to be more immune to noise than the band power features. On the other hand, each PLF feature vector has 37 entries, while the band power features only amount to 14 (the number of considered EEG channels over the motor cortex). Therefore, the PLF features may distinguish the various tasks better simply because they operate on a larger vector. The question remains whether the band power features would produce better results if more EEG channels were used. Nevertheless, the PLF features demonstrate that an approach based on the actual physiology of the production of motor tasks produce better results.

For the band power features it is evident a trend where the imagined tasks produce higher accuracies than the actual tasks. This is probably due to the fact that the trials with actual movement contain more artifacts than the imagined trials, in particular when the subject has to move the legs. This trend is not visible for the PLF features, accentuating the notion that they are less susceptible to noise.

In respect to the alternative hierarchical classifiers, no firm conclusions can be made. There is simply not enough statistical significance in the results. More detailed tests need to be done, comparing the different combinations of motor tasks. Despite this, it is interesting to note that the best classifier (classifier H2) combines, besides the hands, one task from the left side of the body and one task from the right. This was expected, as this type of separation also occurs in the motor cortex, where limbs from the right side of the body are represented on the left brain hemisphere and limbs on the left side of the body are represented on the right hemisphere.

The slight increase in accuracy, from 86.3% for the Hierarchical PLF Classifier with imagined tasks to 89.7% for the H2 classifier is probably due to the reduction of the number of classes, with two less classes to confuse the classifier.

In general terms, the results obtained with PLF features, considering all the classes, are on par with current literature (see, for example, [14] and [29]). Note, however, that these studies use only three or four motor tasks. Comparing with the results obtained for the alternative classifiers, the proposed method provides increased accuracy, although the limited number of test subjects does not allow to ascertain its statistical significance.

6.2.2.A The Thresholding Method

Clearly, the weakest point of this thesis is the thresholding method used to select the time instants when the task is actually being performed. The obtained results are, to say the least, temperamental. The reasons for this are three-fold. First, the thresholding method does not use all the available information, being based on a very reduced subset of features (4 channels for band power features, 2

pairs of channels for the PLF features). The problem here is that the ERD event is expected to occur in a localized region of the brain, although there is some spreading to other areas. This makes the definition of a single threshold that fits all the motor tasks very difficult. A more clever way of doing this would be the use of some kind of Blind Source Separation, attempting to separate the information relative to each type of task. Second, the reference segment, i.e., the first second of the trial, is not long enough to obtain good statistics to compute a baseline to be compared with the rest of the trial. And, thirdly, the threshold does not automatically adapt to the signal within the motor task period. Nevertheless, the method served its purpose without the need to use more complex techniques.

6.2.3 Real-Time BCI

As is obvious from the simulation made in Chapter 5, the methodologies used in this thesis are not quite ready yet for a real-time application. First of all, a better understanding of the PLF features is necessary, with more data on which to test the various approaches. Secondly, some changes would need to be made to the signal processing steps. For instance, the filters could no longer be applied in both directions, as only a limited amount of data is available for processing at a given time. Additionally, all implementations should be as computationally efficient as possible, in order to minimize the response delay of the system. In this respect, the processing window, here chosen with a length of $256\,ms$, should also be smaller, but long enough to provide useful information for the classification. Also, the MATLAB® environment is not the most adequate to implement a system of this type, not being as efficient as other programming languages. In regard to the training of the classifier, an off-line approach, like the one used here, could be implemented in the beginning of the session, although the acquisition conditions change over time and, therefore, it would be necessary to use an adaptive approach, retraining the classifier as new information arrives. Finally, a good feedback system has to be incorporated in order for the user to understand clearly what the BCI is doing. It would be very interesting if we could deliver this information directly to the brain, i.e., a bidirectional BCI, as was done in some recent experiments [30].

6.3 Conclusion

The most important conclusion to extract from the work developed in this thesis is that a BCI system based on the use of PLF features is better than an equivalent system based on power band features, considering the limited data available. Furthermore, the system is capable of distinguishing between seven different motor tasks, which is unusual for this type of approach. Nevertheless, more research is still needed, and much remains to do in order attain the next step, which is to adapt this system into a real-time, fully functional, application.

In a more general view, this thesis accomplishes its goal of being one last test to the knowledge and experience obtained over the entire Biomedical Engineering course at IST, allowing a smooth transition to the work market.

Bibliography

[1] U. Hoffmann, J. Vesin, and T. Ebrahimi, "Recent advances in brain-computer interfaces," in IEEE 9th Workshop on Multimedia Signal Processing, 2007.

[2] T. Ebrahimi, J. Vesin, and G. Garcia, "Brain-computer interface in multimedia communication," IEEE Signal Processing Magazine, vol. 20, pp. 14–24, 2003.

[3] M. Almeida, J. Bioucas-Dias, and R. Vigário, "Source separation of phase-locked subspaces," in Proceedings of the International Conference on Independent Component Analysis and Signal Separation, vol. 5441, 2009, pp. 203–210.

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

[5] E. Schmidt, J. McIntosh, L. Durelli, and M. Bak, "Fine control of operantly conditioned firing patterns of cortical neurons," Experimental Neurology, vol. 61 Issue 2, pp. 349–369, 1978.

[6] L. R. Hochberg, M. D. Serruya, G. M. Friehs, J. A. Mukand, M. Saleh, A. H. Caplan, A. Branner, D. Chen, R. D. Penn, and J. P. Donoghue, "Neuronal ensemble control of prosthetic devices by a human with tetraplegia," Nature, vol. 442, pp. 164–171, 2006.

[7] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robonson, and T. M. Vaughan, "Brain-computer interface technology: A review of the first international meeting," IEEE Transactions on Rehabilitation Engineering, vol. 8, pp. 164–173, 2000.

[8] J. Hawkins and S. Blakeslee, On Intelligence. Owl Books, 2005.

[9] A. Bashashati, M. Fatourechi, and R. K. Ward, "A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals," Journal of Neural Engineering, vol. 4, pp. R32–R57, 2007.

[10] W. Jin, Z. Jiacai, and Y. Li, "An automated detection and correction method of eog artifacts in EEG-based BCI," 2009.

[11] S. Halder, M. Bensch, J. Mellinger, M. Bogdan, A. Kübler, N. Birbaumer, and W. Rosenstiel, "Online artifact removal for brain-computer interfaces using support vector machines and blind source separation," Computational Intelligence and Neuroscience, vol. 2007, p. 82069, 2007.

[12] M. Schröder, M. Bogdan, W. Rosenstiel, T. Hinterberger, and N. Birbaumer, “Automated EEG feature selection for brain-computer interfaces,” in Proceedings of the 1st International IEEE Conference on Neural Engineering, 2003, pp. 626–629.

[13] R. Leeb and G. Pfurtscheller, “Walking through a virtual city by thought,” in Proceedings of the 26th Annual International Conference of the IEEE EMBS, 2004, pp. 4503–4506.

[14] A. R. Murguialday, V. Aggarwal, A. Chatterjee, Y. Cho, R. Rasmussen, B. O’Rourke, S. Acharya, and N. V. Thakor, “Brain-computer interface for a prosthetic hand using local machine control and haptic feedback,” in Proceedings of the 2007 IEEE 10th International Conference on Rehabilitation Robotics, 2007.

[15] G. Edlinger, C. Holzner, C. Guger, C. Groenegress, and M. Slater, “Brain-computer interfaces for goal orientated control of a virtual smart home environment,” in Proceedings of the 4th International IEEE EMBS Conference on Neural Engineering, 2009, pp. 463–465.

[16] B. H. Dobkin, “Brain-computer interface technology as a tool to augment plasticity and outcomes for neurological rehabilitation,” Journal of Physiology, vol. 579.3, pp. 637–642, 2007.

[17] T. M. Vaughan, D. J. McFarland, G. Schalk, W. A. Sarnacki, D. J. Krusienski, E. W. Sellers, and J. R. Wolpaw, “The Wadsworth BCI research and development program: at home with BCI,” IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 14, pp. 229–233, 2006.

[18] [Online]. Available: http://www.neurosky.com/

[19] A. C. Guyton and J. E. Hall, Textbook of Medical Physiology. Elsevier Saunders, 2005.

[20] G. Pfurtscheller and F. H. Lopes da Silva, “Event-related EEG/MEG synchronization and desynchronization: basic principles,” Clinical Neurophysiology, vol. 110, pp. 1842 – 1857, 1999.

[21] S. Halder, D. Agorastos, R. Veit, E. M. Hammer, S. Lee, B. Varkuti, M. Bogdan, W. Rosenstiel, N. Birbaumer, and A. Kübler, “Neural mechanisms of brain-computer interface control,” NeuroImage, vol. 55, pp. 1779 – 1790, 2011.

[22] R. Takalo, H. Hytti, and H. Ihalainen, “Tutorial on univatiate autoregressive spectral analysis,” Journal of Clinical Monitoring and Computing, vol. 19, no. 6, pp. 401–410, 2005.

[23] D. J. Krusienski, D. J. McFarland, and J. R. Wolpaw, “An evaluation of autoregressive spectral estimation model order for brain-computer interface applications,” in Proceedings of the 28th IEEE EMBS Annual International Conference, 2006, pp. 1323–1326.

[24] D. J. McFarland and J. R. Wolpaw, “Sensorimotor rhythm-based brain-computer interface (BCI): model order selection for autoregressive spectral analysis,” Journal of Neural Engineering, vol. 5, pp. 155–162, 2008.

[25] S. Kalitzin, J. Parra, D. N. Velis, and F. H. L. da Silva, “Enhancement of phase clustering in the EEG/MEG gamma frequency band anticipates transitions to paroxysmal epileptiform activity in

epileptic patients with known visual sensitivity," IEEE Transactions on Biomedical Engineering, vol. 49, pp. 1279–1286, 2002.

[26] G. Townsend and Y. Feng, "Using phase information to reveal the nature of event-related desynchronization," Biomedical Signal Processing and Control, vol. 3, pp. 192–202, 2008.

[27] T. M. Mitchell, Machine Learning. McGraw-Hill, 1997.

[28] E. Arbabi, M. B. Shamsollahi, and R. Sameni, "Comparison between effective features used for the bayesian and the SVM classifiers in BCI," in 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, 2005.

[29] B. Blankertz, F. Losch, M. Krauledat, G. Dornhege, G. Curio, and K.-R. Müller, "The berlin brain–computer interface: Accurate performance from first-session in BCI-naïve subjects," IEEE Transactions on Biomedical Engineering, vol. 55, pp. 2452–2462, 2008.

[30] [Online]. Available: http://www.technologyreview.com/biomedicine/38792/