KAUNAS UNIVERSITY OF TECHNOLOGY

IGNAS MARTIŠIUS

DATA ACQUISITION AND SIGNAL PROCESSING METHODS FOR BRAIN – COMPUTER INTERFACES

Doctoral dissertation
Technological sciences, informatics engineering (07T)

2016, Kaunas

UDK 004.5+616.8-009.1] (043.3)

Doctoral dissertation was prepared in Kaunas University of Technology, Faculty of Informatics, Department of Software Engineering, during the period of 2011 – 2016.

Scientific supervisor:

Prof. dr. Robertas DAMAŠEVIČIUS (Kaunas University of Technology, technology sciences, informatics engineering – 07T)

Doctoral dissertation has been published in:
http://ktu.edu

Editor:

Giedrė Drėgvaitė (Kaunas University of Technology)

ISBN 978-609-02-1197-7

KAUNO TECHNOLOGIJOS UNIVERSITETAS

IGNAS MARTIŠIUS

DUOMENŲ SURINKIMO IR SIGNALŲ APDOROJIMO ALGORITMAI SMEGENŲ – KOMPIUTERIO SĄSAJAI

Daktaro disertacija
Technologijos mokslai, Informatikos inžinerija (07T)

2016, Kaunas

UDK 004.5+616.8-009.1] (043.3)

Disertacija rengta 2011 – 2016 metais Kauno technologijos universiteto informatikos fakultete programų inžinerijos katedroje.

Mokslinis vadovas:

Prof. dr. Robertas DAMAŠEVIČIUS (Kauno technologijos universitetas, technologijos mokslai, informatikos inžinerija – 07T).

Doctoral dissertation has been published in:
http://ktu.edu

Redagavo:

Giedrė Drėgvaitė (Kauno technologijos universitetas)

ISBN 978-609-02-1197-7

Contents

List of Figures

List of Tables

GLOSSARY

AAR	Adaptive autoregressive
AC	Alternating current
AEP	Auditory-evoked potential
AI	Artificial intelligence
AR	Autoregressive
ANN	Artificial neural network
ALS	Amyotrophic lateral sclerosis
AUC	Area under the curve
BBCI	Berlin brain-computer interface
BCI	Brain-Computer-Interface
BLDA	Bayesian linear discriminant analysis
BR	Bayesian regularization training function
CAR	Common average reference
CGF	Fletcher-Powell conjugate gradient backpropagation training function
CSP	Common spatial patterns
DCT	Discrete cosine transform
DEO	Discrete energy operator
DFT	Direct Fourier transform
DSP	Digital signal processing
ECG	Electrocardiogram
EEG	Electroencephalogram
EP	Evoked potential
EOG	Electrooculogramm
ERD	Event related desynchronization
ERS	Event related synchronization
ERP	Event-related potentials
FFT	Fast Fourier transform
FIR	Finite impulse response
FLDA	Fisher linear discriminant analysis
fMRI	Functional magnetic resonance imaging
FPR	False positive rate
HCI	Human-Computer Interaction
HMPO	Homogenous multivariate polynomial operator
HOEO	Higher order energy operator
ICA	Individual component analysis
IFSECN	International Federation of Societies for Electroencephalography and Clinical Neurophysiology
IIR	Finite impulse response
LDA	Linear discriminant analysis

LIS	Locked-in syndrome
LM	Levenberg-Marquardt training function
MEG	Magnetoencephalography
MLP	Multi layer perceptron
MMC	Matthews correlation coefficient
N100 (N1)	The first negative component of an event related potential, appearing 100 ms after stimulus presentation
P300 (P3)	Event related potential component elicited in the process of decision making.
PCA	Principal component analysis
PSD	Power spectral density
RAM	Random access memory
RBF	Radial basis functions
RFLDA	Regularized Fisher linear discriminant analysis
ROC	Receiver operating characteristic
SCP	Slow cortical potential
SL	Surface Laplacian
SSAEP	Steady-state auditory-evoked potential
SSEP	Steady state evoked potential
SSVEP	Steady-state visually-evoked potential
STFT	Short-time Fourier transform
SVM	Support vector machine
TKEO	Teager-Kaiser energy operator
VEP	Visually-evoked potential
VTEO	Variable length Teager-Kaiser energy operator
WAT	Wave atom transform

1. INTRODUCTION

1.1. Relevance of the work

Computer Human-computer interaction has been an important research concept since the invention of the computer. Methods of computer interfacing have progressed from perforated cards, keyboards and mice to touchscreen and speech recognition. HCI devices allow humans to interface with computers for the purposes of data entry, control or communication. Most of the efforts over the years have been dedicated to the design of user-friendly, efficient and ergonomic systems to produce a faster and more comfortable means of communication. Interfaces based on voice recognition, gesture recognition, physical movement and other technologies have received enormous research attention over the years and successful examples of these technologies are being produced commercially.

Recently, a radical and novel approach of computer interfaces has received a lot of scientific interest. Since the first experiments of electroencephalography on humans in 1929, the electroencephalogram (EEG) of the human brain has been used mainly to evaluate neurological disorders in the clinical environment and to investigate brain functions in the laboratory. An idea that brain activity could be used as a communication channel has gradually emerged.

The possibility of recognizing a single message or command, considering the complexity, distortion, and variability of brain signals appeared to be extremely remote. Yet EEG demonstrates direct correlations with user intentions, thereby enabling a direct brain-computer communication channel.

Direct brain-computer communication requires high computational capacity to analyze brain signals in detail and in real-time, and until recently the requisite technology either did not exist or was extremely expensive.

The continuing development of computer hardware and software now supports highly sophisticated online analysis of many signal channels at high speed. Also, greatly increased social recognition of the needs and potential contributions of people with severe neuromuscular disorders such as spinal cord injury has generated clinical, scientific, and commercial interest in better communication and control technology.

An interdisciplinary field of research has been created to offer direct HCI via signals, generated by the brain itself. Brain-Computer Interface (BCI) technology, as it is known, is a communication channel that enables users to control devices and applications without the use of muscles. The development of cognitive neuroscience field has been instigated by recent advances in brain imaging technologies such as electroencephalography, magnetoencephalography and functional magnetic resonance imaging.

The growing field of BCI research is relatively new. The first BCI prototype was created by Dr. Vidal in 1973 [Vidal, 1973]. This system was intended

to be used as a promising communication channel for persons with severe disabilities, such as paralysis, amyotrophic lateral sclerosis (ALS), brain stroke or cerebral paralysis [Kübler et al., 2001].

It is estimated that about one million people worldwide suffer from the so-called locked-in syndrome, but despite their paralysis, retain full brain function, which could be employed for communication.

BCI research has been successfully used not only for helping the disabled, but also as being an additional data input channel for healthy people. It can be exploited as an extra channel in game control, augmented reality applications, household device control, fatigue and stress monitoring and other applications.

BCI design represents a new frontier in science and technology that requires multidisciplinary skills from fields such as neuroscience, engineering, computer science, psychology and clinical rehabilitation. Despite recent developments, there are numerous obstacles to building a usable and effective BCI system. The biggest challenges are related to accuracy, speed, price and usability. Current BCI systems are inaccurate and have a low information transfer rate. This means that the user needs a long period of time in order to send commands to the device that is being controlled.

Another problem is the high cost of EEG equipment, such as an EEG cap and amplifiers [Holz et al., 2013]. Systems with a high sensor count take a long time to prepare for use and are uncomfortable. Due to these limitations, no BCI system has become commercially successful to this date. If a disabled person can move their eyes or even one muscle in a controlled way, the interfaces based on eye-gaze or muscle movement (electromyogramm - EMG) switch technology are more efficient than any of the BCIs that exist today.

A sound knowledge of the data acquisition process, EEG waveform characteristics, signal processing methodologies for feature extraction and classification is a prerequisite before attempting to design and implement a functional BCI system. These research points have been highlighted by the BCI development community as being both important and necessary, for further BCI development [Guger et al., 2015, McFarland et al., 2006, Kubler et al., 2006, Krusienski et al., 2011].

1.2. Object of the work

The object of this research is to study EEG signal processing and classification techniques in order to design a Brain-Computer Interface system architecture with a EEG signal as the input modality.

1.3. Aim of the work

This thesis focuses on BCI data acquisition, signal processing and classification techniques for BCI. The aim of this work is to develop new signal denoising and feature extraction algorithms, as well as integrating these algorithms into a BCI system design, based on the Emotiv EPOC neuroheadset

device. The system is to be both end-user friendly, cost effective, wearable, wireless and provide reasonable usability and quality of control for a domestic BCI. It is envisioned to design and implement a non-invasive, affordable and simple to use solution.

1.4. Tasks of the work

The main task of this study is to investigate and design feature extraction and classification algorithms and architectures. The main tasks of this work are:

1. Investigate advances in the field of neurotechnology, and the state-of-the-art BCI system structure;
2. Proposing a classification scheme for BCI systems;
3. Designing a signal denoising technique for EEG;
4. Proposing feature extraction algorithms for BCI;
5. Producing a practical implementation of the proposed algorithms by developing a BCI system architecture.

1.5. Scientific novelty

The work achieved these novel results:

1. A novel Class-adaptive signal denoising scheme has been proposed for use with EEG data.
2. In order to improve feature extraction, the Wave Atom transform has been applied in EEG data processing. By testing this method with standard datasets and comparing results, we have shown that this method can be successfully used for feature extraction and dataset feature reduction in BCI systems.
3. A novel non-linear operator, named Homogeneous Multivariate Polynomial Operator (HMPO) was proposed for signal feature extraction . This operator can be used for developing new EEG processing algorithms.
4. Since BCI systems are subject-dependent, a time-bound voted perceptron artificial neural network training algorithm, bound by real-time constraints has been developed. This algorithm is suited for real time BCI systems, which require constant classifier retraining.
5. We propose a method for fractional time-delay embedding, by selection of positive and negative trajectories from the phase space, distance adaptive sampling and reconstruction of the high-dimensional phase space of the EEG signal.

1.6. Practical value

Systems used today are bulky, wired and laboratory-oriented. They are uncomfortable and inconvenient for users, impending their daily tasks. Most signal processing is performed off-line, hindering real-life applications.

BCI systems require correct classification of signals interpreted from the brain for useful operation. The developed algorithms allow for an improvement in classification accuracy, thereby making them useful in the development of BCI systems.

Such a system would be extremely useful for disabled patients, provided it was portable, wearable and capable of monitoring signals remotely via a wireless transmission protocol. First a wireless system would reduce installation complexity, provide the user with more freedom of movement, so they could perform routine tasks in real-world environments. This would allow for a much wider application of BCI systems [Lin et al., 2010].

This thesis presents the fundamental knowledge in EEG based BCI development, signal processing and machine learning-based classification. It also presents a state-of-the-art review of BCI research and then describes a system implemented by the author.

1.7. Thesis statements

1. Since the EEG signal is non-linear and non-stationary it requires complex feature extraction methods. The Wave Atom transform was used to reduce EEG feature complexity, thereby achieving good classification results.

2. Nonlinear operators were considered for EEG signal processing. A novel Homogenous Multivariate Polynomial Operator (HMPO) was proposed, achieving classification accuracy of 94% with a standard dataset.

3. Since BCI classifiers require constant retraining, a time-bound Voted Perceptron Aritifical Neural Network training algorithm was proposed, achieving good classification results with 3 different datasets, with improved training time, suitable for real-time BCI systems.

4. An analysis of the BCI domain reveals that although consumer-grade devices are incapable of quality signal acquisition, they can be implemented in a BCI architecture. A 3-class self-paced BCI system, based on SSVEP and the Emotiv EPOC headset achieved good accuracy and usability results.

1.8. Scientific approval

All of the results presented in the thesis are original and correspond to a total of 15 publications. 4 internationally referred "ISI Web of Science" scientific

journal publications with a citing index, 5 publications in "ISI Web of Science" journals without a citing index and 6 conference preceedings publications in the fields of informatics, electronics and machine learning journals. The remaining are experimental setups, small unpublished observations or well known facts.

The experimental results were presented in and discussed in 4 international conferences:

1. 11th International conference on Artificial intelligence and soft computing, ICAISC 2012;

2. 18th International conference on information and software technologies, ICIST 2012;

3. 18th International conference Electronics 2014, Palanga;

4. 15th International PhD Workshop OWD 2013, Wisla.

1.9. Thesis organization

The thesis consists of 5 chapters: Chapter 1 is an introduction, providing a short summary of the works novelty, aims and objectives.

Chapter 2 performs a thorough review of electroencephalography (EEG), a brain imaging technology based on the electrophysiological activity within the brain. The chapter describes in detail the origin, functional behavior, acquisition, characterization, taxonomy and applications of EEG signals. The purpose of this chapter is to familiarize the reader with terminology and EEG characteristics that will be exploited and referred to in later chapters.

Chapter 3 begins by introducing the idea and purpose of a BCI. The essential components of a BCI framework are described and some of the signal-processing methodologies behind them are reviewed in detail. A large portion of this chapter is a review of BCI technology and a description of the various approaches by different BCI research groups worldwide. Different performance metrics are reviewed. The chapter concludes by reviewing necessary standardized performance metrics and discusses the challenges for future progression of this technology.

Chapter 4 describes the offline datasets used for direct result comparison with other authors. It introduces proposed novel methods for BCI applications, including signal filtering and classification algorithms. It also presents an architecture of the cost effective BCI system, based on Emotiv EPOC hardware (described in detail in Chapter 4.7.1) and novel algorithms. The chapter describes the brain activity associated with visual stimulation and the methodologies that are exploited in this study to offer control, presents a real-time BCI controlled video application. The real-time deployment of this system and the associated performance results are reviewed. Finally, a discussion reviews the success and future work of this type of BCI implementation.

Chapter 5 provides a conclusion on the issues addressed by this research and on the future of BCI technology.

2. BACKGROUNDS OF ELECTROENCEPHALOGRAPY

This chapter serves as an introduction to the electrobiological brain signal acquisition known as electroencephalography (EEG). The neurophysiological and anatomical structure of the human brain is presented to provide terminology and to serve as background information on its structure and function. The chapter focuses on the origin of electrical potentials that are registered by electrodes, thereby creating EEG. The subsequent section highlights the methods, potential complications and equipment necessary for EEG acquisition. Some sections also explain EEG signals from the digital signal processing (DSP) point of view. This will form a basis for later chapters, when particular features of EEG that identify with a curtain event or function are used for the purposes of BCI development.

2.1. Anatomical background

The human brain consists of over 1 trillion (10^{12}) cells, 1 billion (10^{9}) of which are nerve cells, connected to the neural network [Buzsaki, 2006]. Neurons may be interconnected one-to-one, one-to-many and many-to-many, similar to digital logic AND-OR circuitry in computers. Although different brain regions are responsible for quite different mental tasks, they are connected together in an information sharing network. The cerebral cortex is the brain's outer layer of neuronal tissue. It is located directly behind the skull and is 2 to 4 millimeters thick. The cortex is responsible for tasks such as thought, language, consciousness, memory, attention and other higher level mental processes, and can be subdivided into several areas, such as the motor cortex, somatosensory cortex, visual cortex, etc...

In addition, the cerebral cortex may be classified into four lobes, on the basis of topographical conventions: the temporal lobe, occipital lobe, parietal lobe and frontal lobe. Their location is presented in Figure 2.1. Since the architecture of the brain is not uniform and the cortex is functionally organized, the EEG can vary depending on the location of the recording electrodes.

The functional activity of the brain is highly localized. This facilitates the cerebral cortex to be divided into several areas responsible for different brain functions. The areas are depicted in Figure 2.2 and their related functions are described in Table 1.

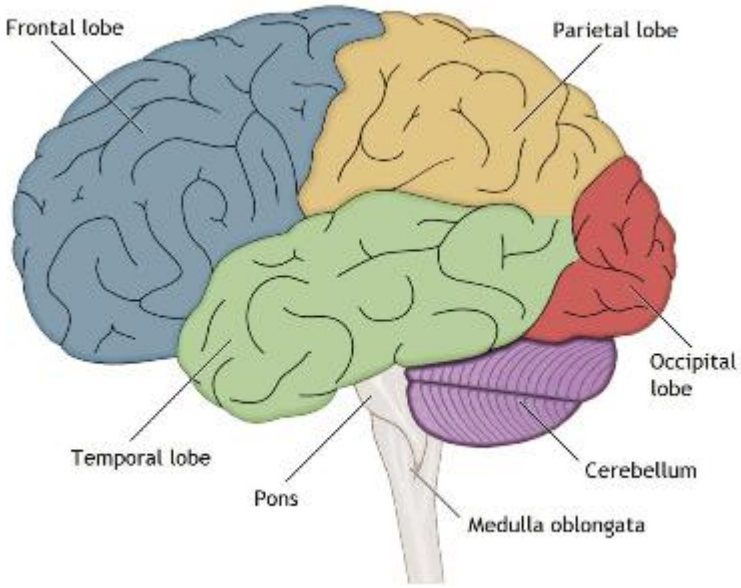


Figure 2.1: Lobes of the human brain [Goldberg, 2002]

Table 1: Cortical areas and their functions

Cortical area	Function
Auditory association area	Complex processing of auditory information
Auditory cortex	Detection of sound quality (loudness, tone)
Broca's area (speech center)	Speech production and articulation
Prefrontal cortex	Problem solving, emotion, complex thought
Premotor cortex	Coordination of complex movement
Primary Motor cortex	Initiation of voluntary movement
Primary somatosensory cortex	Receives tactile information from the body
Sensory association area	Processing of multisensory information
Gustatory area	Processing of taste information
Wernicke's area	Language comprehension
Primary Visual Cortex	Complex processing of visual information

Anatomically the cerebral cortex can be subdivided into two hemispheres:

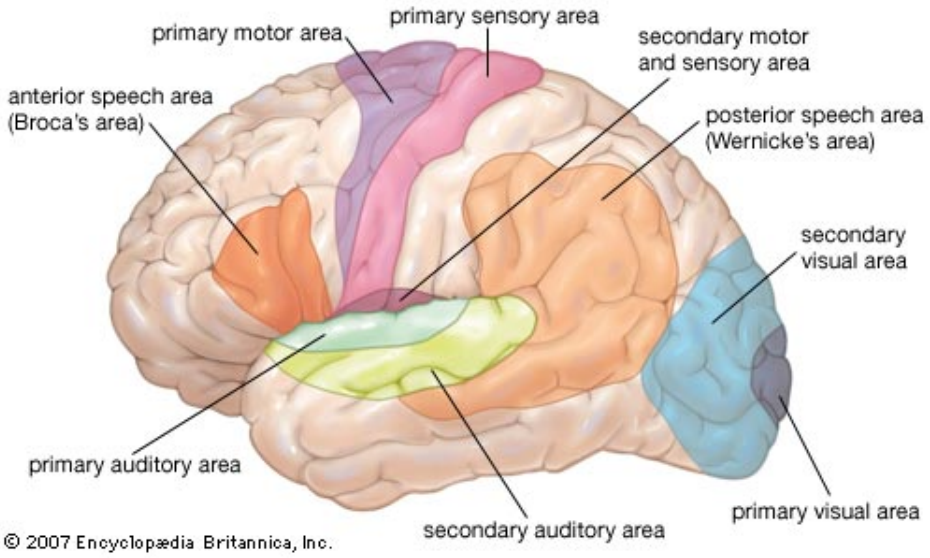


Figure 2.2: Regions of the cerebral cortex [Goldberg, 2002]

the right and the left. They are connected together via the colossal commissure (corpus calossum). The corpus calossum is a stake of about 500 million axon cells, which serve as a information exchange device between the hemispheres [Buzsaki, 2006].

The right hemisphere senses impulses from the left side of the body and is responsible for the movement on the left side. The left hemisphere is similarly connected to the right side of the body.

The main contribution to the electrical EEG signal is generated by extracellular postsynaptic balancing currents between cells and apical dendrites, located in the cerebral cortex. Due to its position, the electrical activity of the cerebral cortex has the greatest influence on EEG recordings. Structures like cell or bone membranes build capacitive resistances, influencing the signal, until finally the signal comes out of the scull and can be picked up by a potential difference measurement device [Buzsaki, 2006].

Another requirement for a measurable signal is the timely synchronization and activation of a whole region of braincells. Only the signal, generated by large groups of neurons is strong enough to be registered by the electrodes, therefore EEG does not register single cell impulses, rather the changing potential of the whole region of neurons.

Electroencephalography is, therefore, a method to locate electrical potentials over different scull areas and map these potentials graphically. Several electrodes are put on the human scalp for this task. A potential difference between two active electrodes on the scalp is measured. This is known as bipolar measurement; Another technique is to measure potential difference between an

active electrode and an electrically inactive reference point (i.e. an electrode on the earlobe) is measured. This is known as reference-point measurement. The measurable EEG activity, used for BCI control is classified into spontaneous and event related potentials. These are further investigated in the following sections.

2.2. Signals for BCI control

2.2.1. Spontaneous potentials

Spontaneous EEG is measured when there is no stimulus presented to the test subject. In healthy subjects the spontaneous EEG is measured during a prolonged time span in which the brain activity changes constant waves into events with higher or lower frequency. Characteristics of different cognitive processes, mental states and activation processes can be observed in spontaneous EEG waves. The appearance of certain frequency bands over a specific brain region can be assigned to a certain mental task. Table 2 shows an overview of different spontaneous EEG frequency bands and their associated functions.

The band range limits associated with the brain rhythms, particularly beta and gamma, can be subject to contradiction and are often further sub-divided into sub-bands that can further distinguish brain processes

Activity with a frequency f, where 30 Hz $\leq f \leq$ 0.5 Hz is often assumed to be of limited clinical utility, although some recent papers have published the existence of cognitive brain process in the beta, gamma and high gamma bands [Palaniappan, 2006], the literature does not clearly state, whether the higher frequency activity (>30 Hz) is of cerebral origin.

EEG rhythms are affected by different actions, thoughts and mental states. For example the planning of a movement can block or attenuate the mu rhythm. The fact that mere thoughts affect the rhythmical activity of the brain can be used as the basis for a BCI system.

2.2.2. Event related potentials

Event related brain potentials (ERP) are different from spontaneous brain activity in the way that they appear while the subject is being stimulated, and are noted by performing extensive analysis of the data. The brain generates not only uninterrupted spontaneous activity, but also reacts to certain external or internal events with a characteristic potential change. On episodic stimulation, event based activity is registered, which is not displayed, if no stimulation is presented.

By presenting the subject with an external stimulus, a specific reaction and specific EEG components are expected to emerge in the ongoing EEG after the stimulus presentation. These event related EEG potentials (known as ERP) are analyzed in the time domain using triggers - timestamps of stimulus

Table 2: Spontaneous EEG frequency bands

Frequency band	Frequency range	Region	Associated activation function
Alpha	8–12 Hz	occipital, parietal	In relaxed wake state with eyes closed
Beta	12–30 Hz	pre-central, frontal	Dominates the EEG when a subject is mentally active, under psychical stress, also in sleep
Delta	1–4 Hz	different	Dominates brain activity in infants and small children. Observed in deep sleep and forced breathing in grown-ups
Gamma	30–80 Hz	frontal	Arises when a subject senses repeated stimulus [Ray and Maunsell, 2011]. Also in understanding of words, meanings, and in sleep.
High Gamma	80–150 Hz	frontal	Stimulus manipulation, attention modulation [Ray and Maunsell, 2011]
Theta	4–8 Hz	frontal, temporal	In a calm relaxed state, deep relaxation, meditation. Very active in small children
Mu	8–13 Hz	frontal-parietal	Recorded over the sensorimotor cortex during waking neural activity [Reilly, 2013].
SCP	0–1 Hz	central	Registered while subject is waiting or preparing for a stimulus. Also observed in relaxation, sleep and coma.

presentation, noted in the EEG.

The subject is presented with a stimulus in constant intervals while his EEG is being recorded. The data encompassing time after the presentation of the stimulus is then analyzed. These induced potentials are called event-related potentials;

A click sound or a flashing light can be used as a stimulus to produce an event-related potential. It is presented repeatedly for several times, sometimes tens or hundreds. The arising event correlated potential with an amplitude of $1/10$ of the spontaneous brain activity, is noise-like and is barely noticeable in

EEG data. After computer analysis of time samples following the stimulus, and by performing averaging on the signals, the evoked potential becomes clearly visible. Most of the oscillations are not of interest, therefore only certain frequencies are measured by selecting a time window of about 100 ms to several seconds. The potential observed has an amplitude of less than 10 µV and a duration of around 0.5 s. It also has a typical form - after a few milliseconds of stimulus presentation, oscillations with a very small amplitude arise. These potential differences are positive or negative changes in brain potential, so one can speak of cortical positivity or cortical negativity [Elbert et al., 1980].

In a typical ERP, first, a small positivity is measured (called P1) [1], followed by a negativity (called N1 or N100, because appearing after approx. 100 ms.), yet again followed by a clear positivity (P3); this is observed after approx 300 ms after the presentation of the stimulus, reaching its peak at about 400 ms, and known as a P300 wave [Gulcar et al., 1998]. An example of an ERP is shown in Figure 2.3.

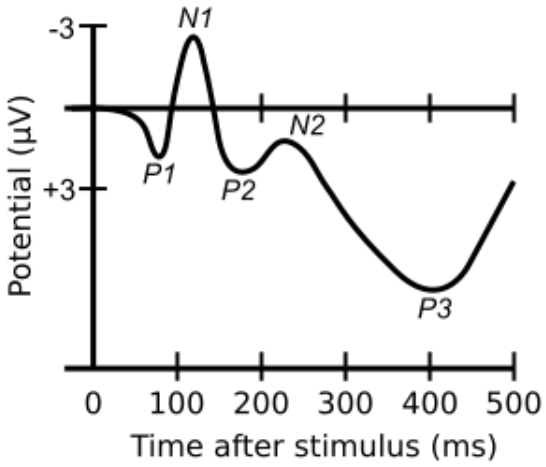


Figure 2.3: Event related potential average of a 100 trials

The P300 and N100 waves are correlative to the stimulus and therefore observed for medical purposes, i.e. in patients with multiple sclerosis, the P300 wave is often longer than in healthy patients. It also serves a purpose for diagnosing other psychological diseases such as schizophrenia, hyperactivity disorders.

Apart from sensory stimulus, event related potentials are evoked by other

[1] It is to be noted, that in the field of neurophysiology, the positive axis is often plotted downwards, and the negative axis-upwards. It is the other way around in most other scientific fields and in some medical literature. There are also authors, who do not agree to this convention. Therefore special attention should be paid to the orientation of the horizontal axis in EEG plots and also positivity and negativity statements

event related actions, such as imaginative or real physical motorical activity, i.e. the movement of arms, legs or the tongue.

2.2.2.1 Evoked-potentials

Evoked potentials (EPs) are a subset of the ERPs that occur in response to or during attention to certain physical stimuli (auditory, visual, somatosensory etc.). They can be considered to result from a reorganization of the phases of the ongoing EEG signals . The EPs can have distinguishable properties related to different stimuli properties, for example, the EEG or Visual Evoked Potential (VEP) over the visual cortex varies at the same frequency as the stimulating light [Raudonis et al., 2008]. Other EPs, such as the Auditory Evoked Potential (AEP) are also used [Rance et al., 1995].

A distinction is made in the literature between a transient EP and a steady state EP (SSEP) based on the stimulation frequency. The former arises when the stimulation frequency is less than 2 Hz. If the stimulus repetition rate is greater than 6 Hz, a periodic response called the SSEP will result.

SSEP are potentials that register when the subject is presented with a periodic stimulus, such as a flickering light or a modulated sound. SSEP are defined by an increase in signal power in the band, equal to the stimulation frequency or integer multiples of that frequency. The amplitude and phase of the SSEP are highly sensitive to stimulus parameters, such as repetition rate, color contrast or sound tone, modulation depth, and spatial frequency. The SSEP was also found to be strongly dependent on spatial attention, being substantially enlarged in response to a repeated stimulus at an attended versus and unattended location. The increased SSEP amplitudes reflect an enhancement of neural responses to a stimulus that falls within the range of spatial attention. It is this fundamental idea that justifies the use of SSEP as a method to identify the attended target among a group of stimuli with sufficiently different stimulation rates.

There are three main modalities of stimulation:

- **Auditory:** signal tones of a specific frequency or clicks are used as stimuli.
- **Visual:** Stimulus is presented as a light with a specific blinking frequency.
- **Somatosensory:** Stimuli are elicited by electrical stimulation of peripheral nerves.

The sequence of stimulation is arranged into paradigms in order to study the responses to certain tasks: The most videly used are:

- **No-task evoked potentials:** The subjects are relaxed and instructed to perform no task upon stimulus reception.

- **Oddball paradigm:** several different stimuli are presented in a random order. In most trials a "non-target" stimulus is presented, and a "target" stimulus is sometimes mixed in. Subjects are instructed to pay attention to the apearance of the target stimulus.

Since chapter 4.7 focuses on a specific evoked potentials, namely, the Steady State Visualy Evoked Potentials (SSVEP), they are presented here in further detail.

2.2.2.2 Steady State Visually Evoked Potentials

A number of studies have demonstrated an increase in neural activity excited by a visual stimulus when the test subject directs his attention to the region of visual space containing the stimulus. [Bi et al., 2013, Shyu et al., 2013, Lee et al., 2011]

The results show that attention acts as a "spotlight", enhancing the cortical representation of stimuli presented in attended regions of visual space relative to stimuli presented in unattended regions of visual space. Studies show that if two or more stimuli with a varying flicker frequency are presented simultaneously, neural responses are elicited by the flicker, receiving the subjects focus. The response generated by the brain corresponds in frequency to the stimulating frequency, and therefore can be detected using Fourier analysis of EEG. In EEG recordings, these steady-state responses are called Steady State Visually Evoked Potentials (SSVEP) [Regan, 1977].

If the subject directs his attention to one visual field and ignores the others while performing a target detection task, SSVEPs elicited by flicker stimulation in the attended visual field have larger amplitude than SSVEPs elicited by the same stimulus in trials where the other field is attended.

The use of frequency-tagging to study attention has the obvious advantage of easily separating neural responses into different classes. How attention modulates the SSVEP response may depend on various parameters, such as stimulus frequency [Shyu et al., 2013], stimulus spacing [Resalat et al., 2012], color [Singla et al., 2013] and shape [Faller et al., 2010]. It is known that low frequency flickering induces more intensive SSVEP, but might cause the users to feel uncomfortable and easily tired.

2.3. Measurement of EEG

2.3.1. Electrodes and electrode placement

In the scalp recorded EEG, the neuronal electrical activity is recorded noninvasively, typically using small metal plate electrodes. An electrode is a conductive plate that picks up electrical activity of the medium that it contacts with. In the case of EEG, electrodes provide the interface between the skin and the recording machine by transforming the ionic current on the skin to

the electrical current in the electrode. Conductive electrolyte media ensures an electrical contact by lowering the contact impedance at the electrode-skin interface. Several types of electrodes are available:

- Disposable (gel-free, and pre-gelled types)
- Reusable cup electrodes (gold, silver, stainless steel or tin)
- Electrode caps
- Needle electrodes
- Nasopharyngeal and Sphenoidal electrodes

For EEG recording the so-called non-polarizing silver (Ag) electrodes, coated with silver chloride (AgCl) are used. These electrodes do not cause polarization, and therefore do not interfere with the signal.

While the number of the electrodes used varies from study to study, they are typically placed at specific scalp locations. The voltages, in the order of microvolts (μV), are carefully recorded to avoid interference and digitized so that it can be stored and viewed on the computer. The amplitude of the recorded potentials depends on the intensity of the electrical source, on its distance from the recording electrodes, its spacial orientation, and on the electrical properties of the structures between the source and the recording electrode. The greatest contributions to the scalp recorded signals result from potential changes which:

- occur near the recording electrodes,
- are generated by cortical dipole layers that are orientated towards the recording electrode at a 90 angle to the scalp surface,
- are generated in a large area of tissue, and
- rise and fall at rapid speed (higher frequency).

2.3.1.1 Electrode positioning

Because EEG on different cortical regions can register place-specific properties, it is vital to record all data with regard to electrode placement.

In order to standardize the methodology of applying the electrodes on the skull, in 1949 the International Federation of Societies for Electroencephalography and Clinical Neurophysiology (IFSECN) adopted a system proposed by Jasper [Jasper, 1958] which has now been adopted worldwide and is referred to as the 10-20 electrode placement International standard, shown in Figure 2.4.

The "10" and "20" refer to the fact that the actual distances between adjacent electrodes are either 10 % or 20 % of the total front–back or right–left

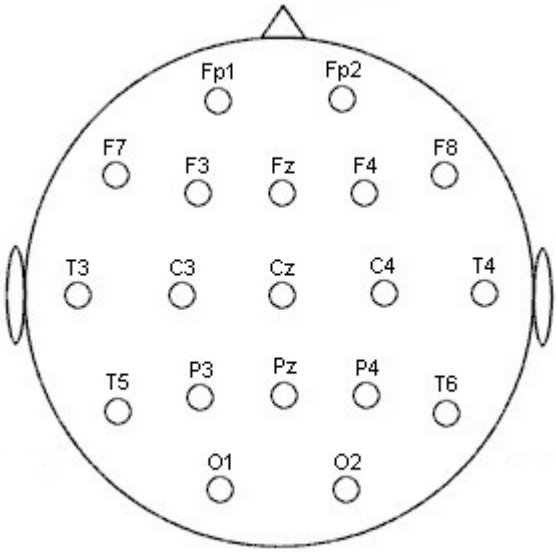


Figure 2.4: The international 10–20 electrode placement system

distance of the skull. This system, consisting of 19 electrodes, standardized physical placement and nomenclature of electrodes on the scalp. This allowed researchers to compare their findings in a more consistent manner. In the system, the head is divided into proportional distances from prominent skull landmarks (nasion, inion, mastoid and preaurical points). Electrode placements are labeled according to adjacent brain regions: F (frontal), C (central), P (parietal), T (temporal), O (occipital). The letters are accompanied by odd numbers for electrodes on the left (ventral) side and even numbers for those on the right (dorsal) side. The letter 'z' instead of a number denotes the midline electrodes. The left and right side is considered by convention from the point of view of the subject. Based on the principles of the 10–20 system, a 10–10 system and a 10-5 system have been introduced in the following years, allowing for more than 300 electrode positions, as extensions to further promote standardization in high-resolution EEG studies (see Figure 2.5).

There are two main forms of electrode coupling, known as reference point and bipolar. Bipolar coupling refers to two differential electrodes placed in the spots of interests. When performing a reference EEG recording, a voltage difference is measured between one point with an electrical activity, and a reference point, on which no electrical activity is supposed to occur. In most cases the ear nipple, the bone behind the ear (Mastoid) or the nose are considered good reference points.

Bipolar measurement has several advantages to reference based measure-

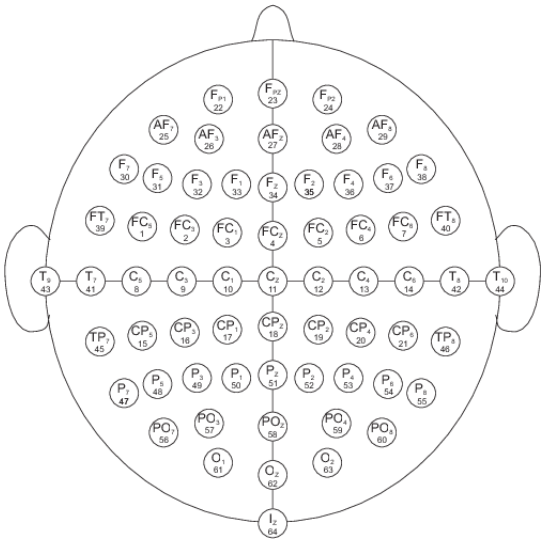


Figure 2.5: The international 10–10 electrode placement system

ment: bipolar couplings are always connected between neighboring electrodes, the known EEG features are visually easier to read in the recordings. Neighboring electrodes are more prone to artifacts, and they are easier to recognize than in reference-point recordings. The lack of a common reference point means the system is more prone to artifacts. The reference point recording technique also has several advantages over bipolar measurement: there is less distortion in signal characteristics such as reduced amplitude, bad polarity, changes in the potential form [Zschokke, 1995].

There is a possibility to filter the signal digitally, after it has been recorded using digital signal processing (DSP) technologies. These mostly rely on Fourier analysis and grand averaging techniques.

2.4. Noise in EEG

Each change in the EEG recording, not directly influenced by the human brain electrical potential, is called an artifact. Noise sources that cause this artificial activity can be classified into sources of biological and sources of technical origin.

Biological artifacts have their roots in the human physiology. They are mostly voltage differences on the scalp caused by extra-cerebral voltage sources, such as muscles. Technical artifacts arise from electrical sources in the pathway from the electrodes to the EEG recording device. They can also be caused by external electromagnetic activity or static electric fields.

The recognition and suppression of artifacts is an important problem in electroencephalography. The origin and characteristics of different artifacts is

also important in recognizing and minimizing their influence. Some artifacts can be similar to real EEG activity. Artifacts, just like EEG features, have to be diagnosed and interpreted. While technical artifacts increase error rates in BCI systems, biological artifacts can cause artificial effects, i.e. eye movement can be a major source of noise in the EEG. If different eye movements are registered in two different experiments, it is hard to recognize, weather EEG activity was also different.

EEG potential differences are in the range of 10–100µV, and therefore have to be registered by sensitive amplifiers. It is obvious that EEG recordings are full of outside electrical interferences – artifacts. Mostly these artifacts differ significantly from the human brain activity and can be easily eliminated. Yet in some cases they are very similar or overlap brain activity.

Artifacts are removed by sophisticated modern equipment, good scalp electrode placement, optimization of the electrical current in the electrodes, filters for known artifact sources, elimination of large artifacts, exceeding 1 mV and EEG variations exceeding 200 µV. This allows for the control of artifacts or their complete removal.

Below is a list of the most common EEG artifacts [Fisch and Spehlmann, 1999]:

2.4.1. Biological artifacts

1. **Lid and eye artifacts** are the most common extra-cerebral noises in EEG recording. In SCP recording, these types of artifacts can be difficult to control. The easiest way to reduce eye interference is gaze fixation: the subject is asked not to blink during the duration of the test. Most subjects find gaze fixation tiring and avoid or ignore it. It is better to measure eye artifacts with an electrooculogram (EOG). This is performed by placing electrodes next or onto the persons eyebrows and excluding samples with high potential changes in these electrodes altogether.

2. **Pulse artifacts** are noise on single electrodes. They arise when an electrode is placed on a pulsating blood vessel. It is a problem of dry, close positioned or defected electrodes. These artifacts are easily differentiated from EEG – the pulse has a uniform pattern. If they, however, are an influence to the reference electrode (i.e. on the mastoid) the human pulse is recorded in all electrodes simultaneously and can be interpreted as Beta waves.

3. **ECG artifacts** are a direct electrical cardiogram intervention into the EEG measurement. They are mostly influencing reference point measurements and are impossible to evade. The ECG is a consequence of a strong muscular dipole, caused by the human heart, which reaches out through the whole human body. It can be 10 times the magnitude of EEG brain

waves. The constant pattern of ECG serves as a feature for its digital removal after EEG recording.

4. **Muscle artifacts** are movements of muscles, mostly close to the scalp area, recorded in the EEG. Therefore these artifacts are recorded over the temporal and frontal scalp areas, where face and neck muscle movements are the main cause of interference. These artifacts are mostly of a shorter duration than EEG and can, therefore be removed by low-pass filters. Grand averaging also helps in removal, because muscle artifacts are random in time and amplitude.

5. **Movement artifacts** arise from subjects body or head movements. The electrode is thereby mechanically moved by the connecting cable. This leads to electrical potential changes that are recorded in the EEG as artifacts. All artifacts arising from muscle movement (i.e. face mimics) breathing and jaw movement artifacts are regarded to as movement artifacts. Movement artifacts, therefore are recorded together, mixed in with the muscle artifacts.

6. **Skin and sweat artifacts** can be caused by changes in skin potential [Lutzenberger et al., 1985]. They appear as low frequency waves (0.2–1 Hz)in the EEG, are irregular and can be correlated to a presented stimulus. They can be detected by a characteristic form of negativity of 2–3.5s, followed by a positive drop. Some cosmetic chemicals and hairsprays can also influence electrode potentials. These are normally eliminated by subject skin preparation in the areas of electrode placement.

   Electrical activity of the sweat glands and resistance changes caused by sweat is a more common source of artifacts. Sweat artifacts are normally recognized by big, synchronous potential changes in several recording channels, mostly in the forehead region. They are partially influenced by laboratory temperature control.

2.4.2. Technical artifacts

1. **Electrode and cable artifacts** can be caused by faulty electrodes or movement of the reference electrode. Reference electrode movement is categorized under biological sources, because it is caused by the movement of the subject (unrest, cough etc.). Electrode artifacts can lead to positive or negative potential changes, but they do not follow any specific rules. Their appearance, however, depends on the type of measurement mode being used. In bipolar connection, they are confined to the electrode pair (single EEG channel) and are observed as a phase shift.

   With a reference point measurement, the noise in the reference leads to a noise in all recording electrodes. These artifacts are mostly distinguished

by sharp potential changes and are easy to detect. Yet if the electrode artifact is a minor potential change, it can be interpreted as a flat Beta wave. Electrode short circuits are a common problem in large electrode count systems and are not noticed most of the time. They are caused by conductive fluid pathways between neighboring electrodes: strong sweating or an excess of conductive gel. Electrode artifacts can also arise from cable connector faults, caused i.e. by the prolonged exposure of the electrode in salt solutions, that causes a chlorination effect on the contact and can lead to false voltages. A light soiling in the contacts can influence resistance. After a period of use, cable cracking can occur.

2. **50 Hz power line frequency and other electromagnetic interferences.** Electromagnetic interferences are artifacts that are caused by alternating current (AC) devices in the vicinity of very sensitive EEG measurement devices. These primarily arise from the 50 Hz (in European countries) alternating current sources such as light and device supply nets. They can lead to artifacts in EEG if there is no proper shielding from other AC sources or a good grounding. Because the frequency of these artifacts is known, they can be eliminated by filtering the signals.

   A distinction needs to be made between two different influences of 50 Hz interferences that arise in EEG recordings. One involves a capacitive or inductive effect of the AC voltage on the subject. The other involves a bad grounding of the subject. A capacitive effect is caused by electrical currents in the laboratory wall and other devices near the subject. Inductive noise is produced by devices with large transformers or electric motors i.e. a ventilator, an electrical field arises in which the subject acts as a antenna.

   Another source of artifacts are high frequency artifacts of over 10 kHz. They are inductively transferred over a large distance, their influence increasing with frequency. Typical sources include radio and microwaves. These do not appear in low frequency EEG, but are converted to a small low frequency voltage in the amplifiers, in a process of demodulation.

3. **Electrostatic interference** is caused by rubbing static electricity. This is caused by certain shoes rubbing against the floor, causing huge electrical potentials on the subject. These can lead to artifacts in EEG, when a ground current between a subject and recording equipment arises.

2.5. BCI illiteracy

A long-standing problem of BCI design that detect EEG patterns is that such paradigms work with varying success among different subjects. In a study done by Florin Popescu, Benjamin Blankertz and Klaus-R. Müller

[Popescu et al., 2008], with a total of 23 untrained users, the group took a close look at how fast the users achieved their best performance during a number of BCI sessions and performance variability among subjects.

Although modern machine learning techniques allow for minimal calibration data size, a strong inter-subject variability exists even after multiple sessions.

The group estimates that about 20 % of subjects do not show strong enough motor-related mu-rhythm variations for effective asynchronous motor-imagery BCI, another 30 % exhibit slow performance (< 20 bits per minute), and up to 50 % exhibit moderate to high performance (20–35 bits/min.).

There is still no clear cause as to why some BCI users exhibit "illiteracy" with BCI systems, and what can be done about it in terms of signal processing and machine learning.

It is clear from several investigations that, BCI illiteracy in a subject appears to depend not so much on the algorithm used but on a property inherent in the subject.

EEG is sensitive to sources in cortical folds, so it might not be able to read motor-imagery activity in some subjects because the particular cortical region involved is tangential to the scalp. An observation consistent with this explanation is that in certain subjects some classes — that is, types of imagined movements—are detectable and others not. Calibration sessions should therefore select subject-specific classes along with frequency bands necessary for feature generation to minimize the illiteracy problem [Popescu et al., 2008].

Several methods have been proposed to detecting solving BCI illiteracy [Vidaurre and Blankertz, 2010, Vuckovic, 2010, Blankertz et al., 2009], yet it remains an open challenge [Krusienski et al., 2011].

2.6. Classification of BCI systems

There is no standard definition of what a BCI is. Different authors define the concept in different ways, thereby causing confusion in the science world and on the general public. Bayliss [Bayliss, 2001] describes a real BCI as an interface which uses only the signals acquired from the human brain. Systems using peripheral nerves and muscles are named Brain-Body actuated. Other authors state that a simple computer keyboard or mouse is a BCI, since they convey the desires of the user to the computer. Both of these definitions have their own flaws, since the Bayliss' definition does not even allow a feedback device, such as a computer monitor to be a part of the BCI. Some authors [Dietrich et al., 2010] claim that he brain contains everything from the network of neurons, up to the abstract higher mental apparatus, the psyche. The term "Brain-Computer Interface" therefore implies an interface between the thoughts of a person and a computer - which is some thing completely different than what can be detected by mere measuring of skull potentials, therefore the term BCI

should not be used at all.

In this thesis the BCI system is considered as a supplementary control channel for computer interfacing by healthy or disabled subjects, as a dependent BCI system, defined by Wolpaw, Birbaumer et al.: a dependent Brain–computer interface (BCI) is a technology for translating brain signals into predefined commands that can be used to communicate with other people or control external devices, without the use of normal brain output pathways. [Wolpaw et al., 2002].

A number of BCI hardware and software devices have been developed since the emergence of the field of neurotechnology. This proposes a need for a classification criteria of different BCI systems. A distinction can be made between evoked potential and self-regulating BCIs, invasive and non-invasive BCIs [Lebedev and Nicolelis, 2006, Wolpaw et al., 2002], neurofeedback based and classification based [Lebedev and Nicolelis, 2006] and dependent and independent [Wolpaw et al., 2002].

2.6.1. Evoked potential and self regulating BCIs

BCI systems can be divided into two distinct classes, depending on subject stimulus presentation. One class is comprised of systems, analyzing evoked potentials – responses of the human brain to a certain stimulus. Farwell und Donchin [Farwell and Donchin, 1988] have measured visually evoked potentials, that were stimulated by blinking letters on the screen. They have shown that the P300 (a positive wave, observed in EEG after 300 ms of stimulus presentation) was higher for the letters the subject wanted to select, than for letters that were not of interest. This component could, therefore, be used for interface control. The systems, developed by Sutter and Tran (1990) [Sutter, 1992], Middendorf, McMillan [Middendorf et al., 2000] belong to this class. The main advantages and disadvantages of these systems is extensively reviewed in [Kübler et al., 2001].

The main thing these systems have in common is that the system presents the stimulus and analyzes the result without the need for active subject participation in the brain wave control process.

The second class of BCI is based on self regulation of certain brain wave frequencies. These systems have a learning process for the subjects, in order to actively learn to control a certain parameter of the brain. The purpose of the learning is to teach the subject the control of a certain EEG parameter, that they can use to control the BCI system for different tasks, i.e. letter, word and sentence writing.

This classification, however, does not assign the systems to distinct classes. When observing the group of invasive BCIs, there is a distinct difference in functional principal and properties. Therefore the field of different measurement method BCIs is further classified.

2.6.2. Invasive and non-invasive BCIs

Invasive and non-invasive interfaces differ in the brain activity measurement technique.

Current invasive methods include single-neuron current measurement, involving direct electrode implantation into the brain tissue [Kennedy et al., 2000, Kennedy et al., 2004, Carmena et al., 2003, Nicolelis et al., 2000] and Electrocorticogram (ECoG) [Lal et al., 2005, Song et al., 2012], which is less invasive. These methods involve surgical procedures to be undertaken for an intercranial insertion of electrodes. The use of these techniques in a BCI is considered to be invasive.

Kennedy et al. implanted a total of 3 patients with an electrode inside the projection area of the hand in the right side of the motor cortex. The electrode contained neurothopical substances, that caused the neuronal axons to attach themselves onto the electrode. After surgery, potentials could be acquired, which could be controlled via neurofeedback training. The patient was trained to lower or raise the electrode potential using auditory and visual feedback. A potential threshold was used for control: patients had to surpass a certain potential threshold, set by the experimenter to achieve cursor or finger movement. The patients controlled a cursor on the 2-D computer screen and a virtual finger, that could flex. One patient achieved successful vertical movement of the cursor. The second patient achieved some control of the virtual finger. The third patient died 76 days after the surgery from an underlying disease without achieving cursor control [Kennedy et al., 2000, Kennedy et al., 2004].

Another study by Song et al. [Song et al., 2012] used 2 patients with implanted ECoG arrays to measure ERP and gamma band oscillations, which were used as features for controlling a letter speller interface. Accuracies reached an average of 80–90 %

The invasive interface has received a lot of criticism: a brain surgery is very risky in itself, so the success of such magnitude is questionable. Animal experiments have shown that potentials from a group of neurons can be used to predict motor activity [Chapin et al., 1999, Nicolelis et al., 2000, Baranauskas, 2014].

Non-invasive BCI methods include EEG [Vidal, 1973, Blankertz et al., 2006a, Muller-Putz et al., 2012], near-infrared spectroscopy (NIRS) [Falk et al., 2011, Yanagisawa et al., 2012, Hu et al., 2010] and functional magnetic resonance imaging (fMRI) [Menon et al., 2013, Lee et al., 2007].

For an BCI interface a physiological parameter which has the shortest latency and is least interfered by breathing and cardiogram artifacts is best selected.

An non-invasive EEG-based BCI uses control of 8–12 Hz Mu rythms, a higher frequency Beta rhythm (20–24 Hz) [McFarland et al., 2006, Park et al., 2013, Chen et al., 2010, Chen et al., 2014] or ERP/ERD poten-

tials. Rhythms are recorded using surface electrodes over the sensomotory cortex and are desynchronized with movement or imaginative movement.

The subjects are given instructions, i.e. imagine a movement of the right hand, and the amplitude of Mu- or Beta rhythm activity measured over the corresponding area of the sensomotory cortex, to control a cursor. Subjects achieved about 90 % accuracy rate after several weeks of training [Blankertz et al., 2006a]. The Mu and Beta rhythm BCI is technically well developed, but has only been tested on healthy subjects and ALS patients in the middle stages of the disease, when they could still move. It remains unclear, whether paralyzed patients would be able to control their Mu rhythm and use this system. The system was also only laboratory-tested, and therefore no long term use, cost or ease-of-use parameters were measured.

2.6.3. Pattern Recognition and Operating Condition based BCIs

This classification is based on the method of control signal extraction from the human brain.

The Operating Conditioning (OC) approach is based on a newly learned ability of the user to control a certain parameter of his brain activity, in a process of *self-regulation*. This learning process is made possible by the feedback of the controlled parameter by the system to the user in a neurofeedback training session. The method is based on user conditioning [Rockstroh, 1989]: on instruction to increase or reduce the feedback parameter S^X, the user produces randomly flickering marker images in the beginning of the experiment. In later training this randomness is replaced by a instruction-induced variation of the parameter R^X, which is sent back via a form of feedback. The concurrence of the instruction and the feedback-signal is calculated and represented as a positive amplification S^R, resulting in a positive consequence C^+ for the user. As training progresses, the possibility of appearance of R^X while showing S^X is increased thereby gradually putting the parameter under user control. Later S^X is replaced by a stimulus S^D, for the amplification reaction R^X. If the user has sufficient control of the parameter, he has the possibility to control the interface. According to Kübler et al. [Kübler et al., 2001] there are three necessary elements required to facilitate the learning procedure in the self-regulation of EEG activity:

1. Real-time feedback of the specific EEG activity,
2. Positive reinforcement of correct behavior,
3. Individual shaping schedule in which progressively more demanding tasks are rewarded.

The OC approach hinges on the effective use of biofeedback in the training scenario for a subject to develop the ability to self-regulate an EEG signal

property. The performances of BCIs based on the OC approach are reasonably successful but the lengthy and arduous training requirement makes it difficult for rapid deployment. Also the number of possible different self-regulation states limits the possible information transfer rates.

This class of BCI has a common trait, that the user must learn to control his brain waves to match a certain pattern. The downside of this setup is that very long training times are required, and since not all subjects' brains are equal, the controlling parameter has to be set up individually for every subject.

An example of an Operating Condition based BCI is the so-called Thought-Translation-Device (TTD), created by Birbaumer et al. [Birbaumer et al., 1999] to enable the user to control a mouse cursor on a computer screen using his slow cortical potentials (SCPs).

Pattern Recognition (PR) BCI systems employ signal processing algorithms for feature extraction and classification, that identify the subjects' brain activity. The user does not need to learn the control of the system, rather the classification algorithm in the system is trained to recognize brain patterns in an activation experiment.

The activity associated with certain mental tasks is highly localized to specific and known regions of the brain. This is used in BCI research to devise suitable experiment paradigms to guide the user to generate distinguishable brain patterns in response to different cognitive mental tasks. The principle of choosing mental tasks is that they produce detectable and different EEG patterns. The Pattern Recognition (PR) approach employs signal processing algorithms for feature extraction and classification that identify and distinguish the brain activity in response to these mental tasks, thus offering a means of control or communication.

The mental tasks used in BCIs up to this date have included motor imagery, visual, arithmetic and baseline tasks. Activation in response to the chosen task should occur close to the cortex so that it can be detected with scalp electrodes.

The classification algorithm is presented with a pattern and a label, signaling an intention of the user to which the pattern belongs. This trains the classification algorithm to learn relations between patterns and labels. This phase is called the training phase. In the latter test phase, the algorithm is presented with a new and unknown pattern and must then decide, which label should be assigned to it.

Using this setup Blankertz et al. [Blankertz et al., 2006a] has used the patterns of imaginary right or left hand or foot movements for training of the classification algorithm. After training, the classifier was presented with new and unknown patterns and could then distinguish, which part of the body the test person was moving in his imagination. The advantage of the latter algorithm, compared to the OC approach is the time it takes for the user to learn

to control the system successfully. The users of TTD had trained for a number of weeks and even months to master the control of their SCPs and gain control of the BCI. The users of the Blankertz et al. BCI were mostly naive users without any BCI control experience, and had undergone the classifier training in less than 20 minutes. This time can be further reduced with modern machine learning algorithms. The OC approach places the biggest requirement on the subject's ability to adapt compared to the PR approach whereby it is up to signal processing and complex classification algorithms to ensure success. The PR approach is technically challenging in ensuring high classification accuracy. It however requires limited classifier training, does not explicitly require feedback, has the potential to offer many control states and can offer a more natural control method.

Decoding brain patterns is a challenge however due to the inter-trial variability of brain activity, the low signal-to-noise ratio (SNR) of the brain response to the mental task with comparison to the background activity, the subtlety between the responses to different tasks and the real-time requirement of the decoding process. It should also be noted that some mental tasks suit some people better than others. It will be shown in the next chapter, that the majority of implemented BCI systems take the PR approach. This method is also the basis of this work.

2.6.4. Dependent and independent BCIs

The discrimination of dependent and independent BCI lies in the need for neuromuscular function to use the BCI. In 1973 dr. Vidal has produced a dependent BCI, with the use of evoked potentials. The user was presented with a matrix of letters on a computer screen. The user focuses his attention on the desired letter in the matrix. Each letter is highlighted in a random order. When a desired letter is highlighted the brain generates a VEP in the visual cortex. The VEP of the focused letter is distinguishable from the VEP of a letter the user did not focus on. Using the VEP impulse time the computer can trace back the letters to select the one which was displayed on the screen at the time of VEP. The user must be able to focus his gaze on a computer screen, using his normal neural pathways and muscles. The system is therefore called a dependent BCI, according to the concept of [Vidal, 1973]. Such systems can not completely perform the bypass function described above. A system which does not use normal neuromuscular pathways completely is considered to be a independent BCI [Wolpaw et al., 2002].

2.7. Synchronous and asynchronous BCIs

A synchronous BCI is a system, which allows user interaction only in specific time periods, which are instructed by the system (cue-pased). The onset of mental activity is therefore externally influenced, and data is analyzed

in a defined time window. Most of todays BCI systems are therefore classified as synchronous. On the other hand, an asynchronous BCI system allows the user to issue commands at any time (self-paced). This allows for a more natural control of the system [Townsend et al., 2004].

This classification can be graphically displayed by using a feature diagram [Damasevicius and Stuikys, 2009], such as shown in Figure 2.6.

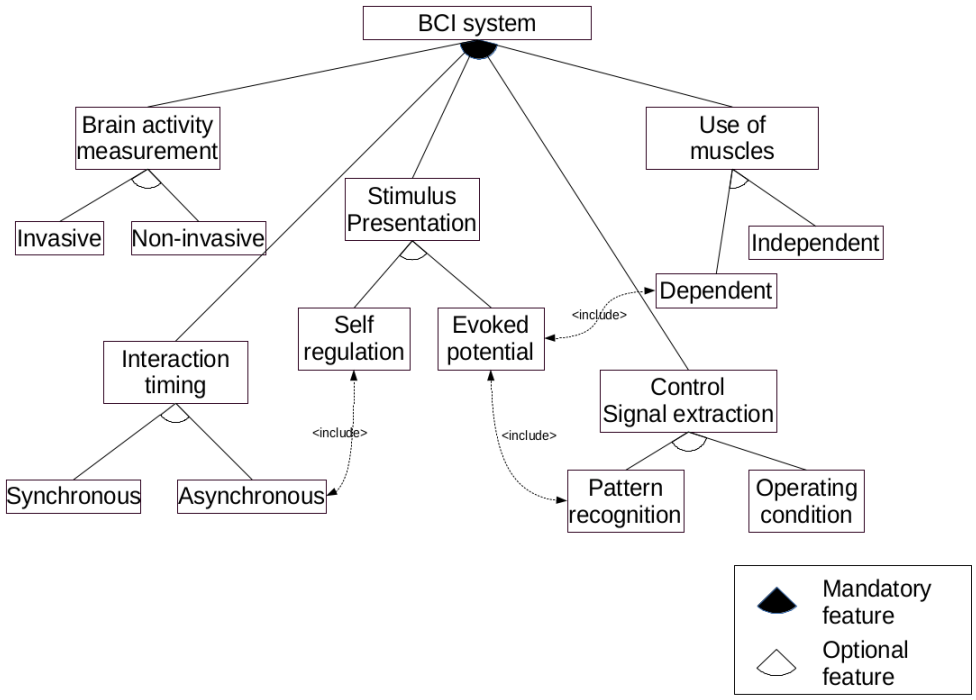


Figure 2.6: Feature diagram of BCI systems

2.8. Applications of BCI

In the past years neuro and biofeedback use for therapeutic disease treatment, and as a communication channel for disabled persons has been extensively researched. The underlying idea is that if a subject receives willful control over a physiological parameter such as brain waves, this can be used for controlling different kinds of appliances, devices and communication systems. A system concept is presented, where subjects that have lost muscle control, and therefore cannot communicate verbally, could use brain wave activity for communication instead. If, however, a patient still has a working muscle, which can be connected to a communication device, this muscle will be faster and more error-prone than a EEG-based system. It would also not require training. If there is no such muscle, or it is unreliable (i.e. because of spasms) or fatigue, an EEG-based communication system could be used instead or as an extra form

of communication.

The target group for brain controlled communication systems are therefore people with a disabled motoric activity, in particular patients with locked-in syndrome (LIS), which still have normal cognitive function. In most LIS patients, vertical eye movements are left intact and are the main communication pathway used in today's systems. In cases of complete-LIS, however, this function is lost, denying the patients eye movement based communication altogether. The life expectancy of LIS patients is 10 years or more [Cappa and Vignolo, 1982].

Another cause of LIS are neurodegenerative diseases such as amyotrophical lateral sclerosis (ALS). ALS is a progressing disease, which blocks voluntary muscle movement. In time it also affects breathing . Only the vegetative muscles of the heart, intestine tract and, for a long time, eye muscles are unaffected. The primary cause of ALS is unknown. There are no known therapy for it. The palliative therapy is to ensure a highest possible quality of life for the patients. A very large factor in life quality is the ability of communication.

The main objective of BCI is to provide severely disabled people with a new communication channel which is not based on the traditional motor output channels. The second category is the multimedia and virtual reality domain. Thus, even if BCI is mainly designed for disabled people, it can also be of interest for healthy persons [Lotte, 2008], for instance by proposing video games based on BCI, cursor, proshtesis, wheelchair control, multimedia and virtual reality applications. This is discussed in more detail in section 3.8. Before this, however, one needs to understand the BCI architecture, which is discussed in the following chapter.

2.9. Technology limitations of BCI

The first challenge for noninvasive BCI technology is the limited information transfer rate (ITR) achievable through EEG, which is about an order of magnitude lower than invasive BCI methods currently provide. The potential benefits of brain implant-based BCI haven't yet proved worth the associated cost and risk in the most disabled patients, let alone healthy users [Popescu et al., 2008].

The biggest problem with most BCIs is low accuracy. Sometimes the output of the system does not match the input. This, of course, can be more or less serious depending on the application. If used for moving the cursor on a computer screen an erroneous output every now and then might be tolerable, but if used for controlling the motion of a physical device, such as a wheelchair, this behavior becomes unacceptable.

Another problem associated with many BCI paradigms is a long input-output delay. Today, the most successful systems work at a transfer rate of less than 30 bits per minute [Dornhege, 2006]. That might be enough to operate a

simple word processor system, but it is too slow to control a wheelchair. Most research today, including the project described in this thesis, therefore focuses on improving the two factors of speed and accuracy of BCI communication.

Even though more and more BCI applications exist, there is still a number of problems BCIs need to overcome to become interesting for the large public. The first problem lies with EEG sensors. Traditional EEG systems like the Biosemi ActiveTwo consist of a cap, a lot of wires and up to 256 EEG electrodes. The high sensor count and wires make such a system impossible to use outside the laboratory, because the setup requires one or several assistants and preparation time. Another drawback is the fact that conductive gel needs to be used for, leaving residue in the user's hair. Furthermore, the system costs thousands of euros. The g.SAHARA system produced by g.tec does not require conductive gel, but still needs wires and an electrode cap.

BCI is often the only input modality in applications which have been developed for research projects. This can be problematic: having to control a cursor continually by means of imagined movement results in a high workload. It would probably be better if BCI was one of multiple modalities used to control an application. Examples of such multimodal applications, or hybrid BCIs, are the 'AlphaWoW'(Alpha-World of Warcraft) [Bos et al., 2010], where brainwaves in the alpha band are combined with keyboard and mouse inputs, and 'Mind the Sheep!'[Gürkök, 2012], where SSVEP is combined with mouse input. Other examples include a touchless system [Zander et al., 2010], which combines eye gaze for cursor control with a BCI for making selections.

The focus of BCI research should shift from reliability to usability and user experience [Kristo et al., 2015]. This is necessary to migrate BCI systems out of the laboratory, into society. Healthy persons can choose from various alternative input modalities. So, for healthy persons to choose BCI, the user experience and usability must be adequate. Most people have never used a BCI and the novelty of this new technology can be a reason for people to decide to use a BCI instead of alternative input modalities, even if BCI is less reliable and slower. However, if the usability is not good, people will choose a different input modality. Due to the fact that the focus in BCI research has mainly been on the reliability, no standardized methods to assess the user experience for BCI yet exist.

The final limitation to mention here is the amount of data, that a BCI can transfer.

The EEG measures a mixture of signals originating in neural brain activity. The two lowest functional layers of the brain are mostly locally oriented. Therefore, the detectable local functional brain areas are responsible for these signals. By observing brain areas, responsible for activating certain muscles, it is possible to measure a corresponding signal (conscious or unconscious) of muscle stimulus.

The higher sublevels (thoughts, emotions), are assembled on the lower levels and exist only in an abstract way. Direct mapping to the physiological media is therefore no longer possible. To measure processes on these higher layers on a physical level would require additional tools to translate the measured low-level information into the higher-level context appropriately - comparable to trying to understand higher levels of information protocols of computer communications using basic electrical measurements. One needs specific 'interpreters' for such operations. The problem is the interpretation of this mixture of measured signals. Hence, for the control of highly complex prostheses, EEG signals are not sufficient, and can never be in the future.

The signals necessary to control the arm, including the consideration of closed loop controls between the brain and the arm, ideally including the fingers and integrated touch sensors, would be too blurred to be the basis for adequate arm movement execution.

2.10. Conclusions

This chapter is aimed at reviewing the anatomical background of BCI technology.

The literature review has shown, that although there are several methods of brain activity measurement, noninvasive EEG is the most convenient, since it is unharmfull to the user and has good temporal resolution.

We have also categorized the signals used to drive the BCI into two categories, which are the evoked potentials, namely P300 and the Steady-State Visually Evoked Potential (SSVEP) paradigms, and spontaneous brain signals, such as Slow cortical potentials (SCP's) and Motor Imagery (MI).

We have then proposed a classification scheme for BCI systems.

Finally, we have reviewed the limitations of BCI systems, that are inherent in the interface.

3. BRAIN COMPUTER INTERFACE DESIGN

3.1. Typical BCI system structure

Designing a BCI system is a multidisciplinary task, involving computer science, signal processing, neurology and physiology. To use a BCI, two phases are required:

1. a training phase, in which:
   (a) the user is trained to willingly control his brain potentials (in the case of operating condition BCI),
   (b) an offline training phase, which calibrates the training algorithm (in the case of pattern recognition BCI),
2. the online phase, in which the BCI system is used for control.

Since this thesis focuses on the pattern recognition approach, we will concentrate on this specific subject.

In the online mode, the BCI system generally performs a six-step process: brain activity measurement, preprocessing, feature extraction, classification, command translation and feedback [Mason and Birch, 2003].

1. Brain activity measurement is the step in which electrodes are used to obtain the user's EEG at specific regions on the scalp, to form input for the BCI system. This step involves determining the number and location of the channels, amplification, analog filtering and A/D conversion. Chanel locations are selected according to the paradigm used and mental task performed.

2. The preprocessing step consists of denoising the recorded brain signal in order to enhance the relevant information inside. Denoising can be performed by channel or artifact rejection, DSP signal filtering methods.

3. Feature extraction is a step to describe the signal by a few relevant, command-related values known as "features".This stage often characterizes the BCI design approach. Features that describe the signal in as few components as possible, are resilient to noise and artifacts have to be identified and used.

4. Classification is a step which assigns a class label to a set of features extracted from the signal. This class label corresponds to the kind of mental state identified. Classification can be performed in various ways ranging from simple thresholding or linear models to complex nonlinear neural network classifiers.

5. Translation into a command is performed by issuing an action, corresponding to the mental state identified i.e. moving the mouse cursor on a computer screen, controlling a speller, moving an automatic wheelchair.

6. The feedback step provides the user with information about the identified mental state. This helps the user to consciously control his brain activity to increase performance.

Such architecture is summarized in Figure 3.1.

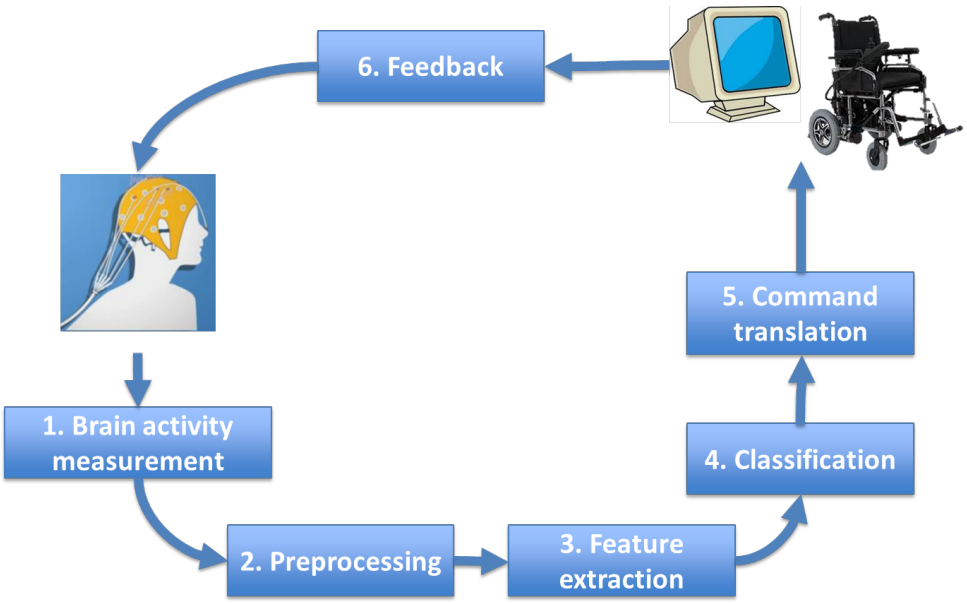


Figure 3.1: General architecture of an online brain-computer interface

Each step is analyzed in detail in the following sections.

These steps define an online BCI. However, as mentioned above, before operating a BCI, a calibration task is performed. This work is generally done offline and is used to calibrate the classification algorithm, selection of optimal features, relevant sensor locations etc. To do this, a training data set must bee recorded by the user. EEG signals are highly subject-specific and therefore BCI systems must be calibrated and adapted to each user individually. The training data set must contain EEG signals recorded while the subject is performing every mental task of interest numerous times. The recorded EEG signals are used as examples to tune the calibration parameters for this subject. The brain patterns of the user can also change in the course of using the BCI. This can happen because of fatigue or changes in the environment. Therefore, a recalibration of the system needs to be performed.

3.2. Preprocessing

Preprocessing involves the preparation of the EEG recordings. It is an important stage that decides the filtering, segmentation and detrending methods used to prepare the EEG data for further stages. Filtering and segmentation (also known as epoching) are used to identify and maximize the information over a certain time or frequency range that is associated with the characteristic brain activity to be elicited.

Most cognitive EEG activity is usually in the range of 0.2 – 40 Hz, thus filtering thereafter reduces noise. A band-pass filter at the electrical mains frequency is typically performed in addition. After filtering, the segmentation of EEG data is performed. This involves splitting the continuous EEG signal into time-locked windows, which usually overlap, or are locked to a stimulus (in the case of synchronous BCIs). Epoching allows for averaging and dramatically simplifies the feature extraction and classification process.

Detrending removes any baseline drift associated with the EEG recordings. This is important to ensure the quasi-stationarity of small EEG segments. The sample rate can be converted to represent the data in as few samples as possible to reduce the computational demands of processing a large number of samples. The sampling rate must be chosen to be at least twice that of the maximum frequency contained in the data (Nyquist rate). A sampling rate of 128 Hz is typically chosen in the literature as it is capable of representing frequencies up to 64 Hz, thus the entire accepted range of EEG activity.

3.2.1. Spatial and Temporal Filters

Brain electrical signals are very weak compared to other biosignals, such as muscle activity or electrocardiogram. They lie in the range of 1–100 µV. For this reason, sensitive amplifiers are used for EEG recording. Artifacts overlaping with the signal of interest dictate the need for signal filters in the amplifier. There are four main types of filters used: highpass - filters that let through only frequencies above a desired threshold; lowpass - filters that let through frequencies below a desired threshold; bandpass - only a desired frequency band is let through; bandstop - restricting a certain frequency band.

A frequency filter can be a time constant, which gives the time period allowed for a signal to be recorded, therefore they are known as temporal filters.

3.2.1.1 Temporal filters

Temporal filters are used in order to restrict analysis to frequency bands in which useful neurophysiological signals are located, i.e., BCI based on sensorimotor rhythms generally band-pass filter the data in the 8-30 Hz frequency band, as this band contains both the μ and β rhythms, i.e., the sensorimotor rhythms [Ramoser et al., 2000]. They are also used to remove other artifacts

such as electrode polarization and power mains noise. This type of filtering is generally achieved using Discrete Fourier Transform (DFT) or using Finite Impulse Response (FIR) or Infinite Impulse Response (IIR) filters.

Direct Fourier Transform (DFT) filtering makes it possible to visualize a signal in the frequency domain, i.e., to see a signal as a sum of oscillations at different frequencies f . Thus, the DFT $S(f)$ of a signal $s(n)$, which is composed of N samples, can be defined as in (1):

$$S(f) = \sum_{n=0}^{N-1} s(n) e^{\frac{-2i\pi fn}{N}}. \tag{1}$$

Filtering a signal using DFT simply consists of removing all coefficients of the DFT which do not correspond to targeted frequencies, and then to transform the signal back into the time domain, by using the inverse DFT, as in (2):

$$S(n) = \frac{1}{N} \sum_{k=0}^{N-1} S(k) e^{\frac{2i\pi kn}{N}}. \tag{2}$$

DFT filtering can be used online and in real-time, thanks to the efficient and popular DFT implementation known as the Fast Fourier Transform (FFT) [Smith, 1999]. Filters can be further optimized to reduce complexity. We describe this method in detail in [Martišius et al., 2011]. The FFT coefficients can also be classified directly, without converting them back to the time domain. Similar methods have been successfully used in [Birvinskas et al., 2012, Birvinskas et al., 2013, Martišius et al., 2013a].

FIR filters are linear filters which make use of the M last samples of a raw signal $s(n)$ in order to determine the filtered signal $\hat{s}(n)$, as shown in (3):

$$\hat{s}(n) = \sum_{k=0}^{M} a_k s(n-k), \tag{3}$$

where the a_k are the filter coefficients. Their values depend on the desired filter. FIR filters are know to have excellent performances in the frequency domain.

As FIR filters, IIR filters are linear filters. On the other hand, IIR filters are recursive filters, which means that, in addition to the M last samples, they make use of the outputs of the P last filterings, as in (4)

$$\hat{s}(n) = \sum_{k=0}^{M} a_k s(n-k) + \sum_{k=1}^{P} b_k \hat{s}(n-k). \tag{4}$$

IIR filters can perform filtering with a much smaller number of coefficients than FIR filters. However, their performance in the frequency domain is slightly reduced [Smith, 1999].

3.2.1.2 Spatial filters

Spatial filters are used to isolate the relevant spatial information embedded in the signals. This is achieved by selecting or weighting the contributions from different electrodes [McFarland et al., 1997]. The most simple spatial filter is created by selecting electrodes over scalp ares, which are known to have relevant brain signals in the current paradigm. Other electrodes are not of interest. The latter electrodes are likely to measure mostly noise or background activity not relevant for the targeted BCI. They can also be used for artifact rejection. As an example, when using a BCI based on hand motor imagery, the neurophysiological signals of interest are localized over the motor or sensorimotor cortex areas. Thus, the main interest is to focus on electrodes C3 and C4, which are located over the left and right motor cortex respectively (see Figure 2.2). Similarly, for BCI based on SSVEP, the most relevant electrodes are the electrodes O1 and O2, which are located over the visual areas [Lalor et al., 2005b]. Electrodes over the eyes, such as Fp1 and Fp2 are commonly used to reject blinking artifacts - trials which have high amplitude values in these channels are discarded from the data.

Other simple and popular spatial filters are the Common Average Reference (CAR) and the Surface Laplacian (SL) filters [McFarland et al., 1997]. These two filters make it possible to reduce the background activity. The CAR filter is obtained by (5).

$$\hat{V}_i = V_i - \frac{1}{N_e} \sum_{j=0}^{N_e} V_j, \tag{5}$$

where $\hat{V}_i$ and V_i are the i^{th} electrode potential, after and before filtering respectively, and N_e is the number of electrodes used. Thus, with the CAR filter, each electrode is re-referenced according the average potential over all electrodes. The SL filter can be obtained as in (6).

$$\hat{V}_i = V_i - \frac{1}{4} \sum_{j \in \Omega_i^4} V_j, \tag{6}$$

where Ω_i^4 is the set of the 4 neighboring electrodes of electrode i. Thus, this filter emphasizes localized activity and can reduce diffuse spatial activity. It is represented graphically in Figure 3.2.

Naturally other more complex preprocessing methods are being proposed and used, such as the common spatial pattern filter.

3.2.1.3 Common Spatial Patterns

The Common Spatial Pattern (CSP) method is based on EEG signal decomposition into spatial patterns [Ramoser et al., 2000, Dornhege et al., 2004,

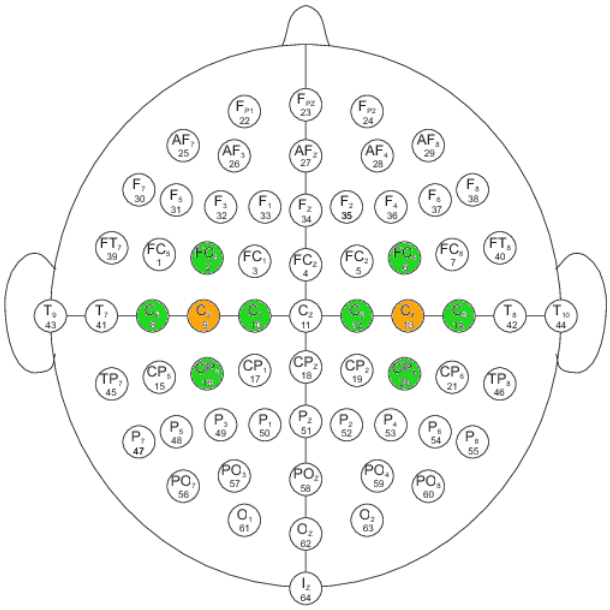


Figure 3.2: Example of a Laplacian filter on channels C3 and C4. $\widehat{C3} = 4 \times C3 - FC3 - C5 - C1 - CP3$

Wu et al., 2005]. These patterns are selected in order to maximize the differences between the classes involved once the data have been projected onto these patterns. CSP is used for feature extraction. It learns spatial filters to combine the existing channels into new virtual channels that are more discriminative than the original ones. Determining these patterns is performed by a joint diagonalization of the covariance matrices of the EEG signals from each class [Ramoser et al., 2000, Dornhege et al., 2004]. The specific process of this algorithm is expressed as described in [Bai et al., 2014]:

1. Calculate the normalized spatial covariance of EEG signals. The two kinds of raw EEG data can be expressed by the $N \times T$ matrices: X_1, X_2, where N is the number of channels, T is the number of samples/channel. The normalized spatial covariance of EEG is obtained by (7)

$$R_i = \frac{X_i X_i^T}{trace(X_i X_i^T)}, i = 1, 2 \,, \tag{7}$$

   where $trace(x)$ is the sum of the diagonal elements of matrix x.

2. Employ the eigenvalue decomposition on the sum of the spatial covariance. It can be given as (8):

$$R = R_1 + R_2 = U_0 \Lambda U_0^T, \tag{8}$$

where U_0 is the matrix of eigenvectors and Λ is the diagonal matrix of eigenvalues.

3. Obtain the whitening matrix and whiten the covariance. The whitening transformation is defined in (9):

$$W = \Lambda^{-1/2} U_0^T, \tag{9}$$

The covariance can be whitened as in (10) and (11):

$$S_1 = W R_1 W^T, \tag{10}$$

$$S_2 = W R_2 W^T. \tag{11}$$

Since S_1 and S_2 share common eigenvectors, that is if $S_1 = U \Lambda_1 U^T$, then $S_2 = U \Lambda_2 U^T$ and $\Lambda_1 + \Lambda_2 = I$, where I is the identity matrix. Since the sum of two corresponding eigenvalues is always one, the eigenvector with a largest eigenvalue for S_1 has the smallest eigenvalue for S_2 and vice versa. Thus, this property makes U useful for classification of two classes.

4. Aquire the projection matrix and decompose a trial. Here, the $2m$ eigenvectors corresponding to the m largest eigenvalues in Λ_1 and Λ_2 respectively are chosen sequentially to form a new matrix U' that are optimal for discrimination of the two classes. The projection matrix can be simplified by $SF = (U')^T W$. The decomposition of a trial X_i is given as in (12):

$$Z_i = SF \cdot X_i, \tag{12}$$

The variance of Z_i can be used as a feature for classification.

Such filters have proved to be very efficient, especially during BCI competitions [Chen et al., 2014, Feng et al., 2013, Li and Guan, 2006]. Currently, CSP are used in the design of numerous BCI systems [Ramoser et al., 2000, Wu et al., 2005, Blanchard and Blankertz, 2004, Congedo et al., 2006, Chen et al., 2014, Wu et al., 2014, Song et al., 2013].

3.3. Feature extraction

Measuring brain activity through EEG leads to the acquisition of a large amount of data. Indeed, EEG signals are generally recorded with up to 256 sensors and with a sampling frequency of 100 Hz – 1 kHz. In order to obtain the best possible performance, one needs a smaller number of values for processing, which describe some relevant properties of the signals. These values are known as "features". Features can be, for instance, the power of the EEG signals

in different frequency bands. Features are generally aggregated into a vector known as a "feature vector". Thus, feature extraction can be defined as an operation which transforms one or several signals into a feature vector.

Identifying and extracting good features from signals is a crucial step in the design of BCI systems. If the features extracted from EEG are not relevant and do not describe the signal well enough, the classification algorithm which will use these features will have trouble classifying the mental states of the user, the correct recognition rates of mental states will be low, in which case the use of the interface would be impossible or inconvenient. Thus, even if it is sometimes possible to use raw signals as the input of the classification algorithm, it is recommended to select and extract good features in order to maximize the performances of the system by making task of the subsequent classification algorithm easier [Verikas and Gelžinis, 2008]. Therefore it is often the case that choosing a good preprocessing and feature extraction method has more impact on the final performances than the choice of a good classification algorithm [Pfurtscheller et al., 1993].

Numerous feature extraction techniques have been studied and proposed for BCI. These techniques can be divided in three main groups, which are:

1. Methods that exploit the temporal information embedded in the signals [Schloegl et al., 1997, Penny and Roberts, 1999, Anderson et al., 1998, Kaper et al., 2004, Geethanjali et al., 2012];

2. Methods that exploit the frequency information [Pfurtscheller and Neuper, 2001, Palaniappan, 2005, del R.Millan et al., 2000, Birvinskas et al., 2013, Birvinskas et al., 2012, Martišius et al., 2013a];

3. Hybrid methods, based on time-frequency representations, which exploit both the temporal and frequency information [Fatourechi et al., 2005, Bostanov, 2004, Boashash et al., 2012].

4. Methods that exploit the spatial information, i.e., inverse solutions. The spatial information is more often used to perform a spatial filtering before extracting features based on the temporal and/or frequential information. [Yitembe et al., 2011, He and Musha, 1988].

3.3.1. Temporal methods

Temporal methods use the temporal variations of the signal as a feature. These methods are particularly adapted to describe neurophysiological signals with a precise and specific time signature, such as the P300 [Koo et al., 2014, Khan et al., 2009, Zhang et al., 2007] or ERD, notably those triggered by motor imagery [Blankertz et al., 2006a]. Temporal feature extraction methods include the raw EEG amplitude, auto-regressive parameters or Hjorth parameters.

3.3.1.1 Signal amplitude

The most simple (but efficient) temporal information that could be extracted is the EEG signal amplitude change in time. Thus, the raw amplitudes of the signals from the different electrodes, possibly preprocessed, are simply concatenated into a feature vector before being passed as input to a classification algorithm. In such a case, the amount of data used is generally reduced by preprocessing methods such as spatial filtering or subsampling. This kind of feature extraction is one of the most used for the classification of P300 [Rakotomamonjy et al., 2005, Hoffmann et al., 2005].

3.3.1.2 Autoregressive parameters

Autoregressive (AR) methods assume that a signal $X(t)$, measured at time t, can be modeled as a weighted sum of its own previous values, to which a Gaussian white noise term E_t can be added, as in (13):

$$X(t) = c + \sum_{i=1}^{p} a_i X_{t-i} + E_t, \tag{13}$$

where the weights a_i are parameters of the model, which are generally used as features for BCI [Wang et al., 2010, Krusienski et al., 2006], c is a constant and p is the model order. Several variants of autoregressive parameters have also been used such as multivariate AR parameters [Cheung and Van Veen, 2008, Jamoos et al., 2011], or Adaptive AR (AAR) parameters [Pfurtscheller et al., 1998, Giannakakis and Nikita, 2008]. AAR parameters assume that the weights a_i can vary over time, and are the most used variant of AR parameters.

3.3.1.3 Hjorth parameters

Hjorth parameters describe the temporal dynamics of a signal $X(t)$, by using three measures – activity (m_0), described by (14) , mobility (m_2), described by (15) and complexity (m_4), described by (16) [Hjorth, 1970].

$$m_0 = \int_{-\infty}^{\infty} S(w)dw = \frac{1}{T}\int_{t-T}^{t} f^2(t)dt, \tag{14}$$

$$m_2 = \int_{-\infty}^{\infty} w^2 S(w)dw = \frac{1}{T}\int_{t-T}^{t} \left(\frac{df}{dt}\right)^2 dt, \tag{15}$$

$$m_4 = \int_{-\infty}^{\infty} w^4 S(w)dw = \frac{1}{T}\int_{t-T}^{t} \left(\frac{d^2 f}{dt^2}\right)^2 dt, \tag{16}$$

where m_n is the spectral moment at order n, $S(w)$ the power density spectrum and $f(t)$ the signal. Such features are mainly used for the classification of motor imagery [Vourkas et al., 2000, Pal et al., 2014].

3.3.2. Frequency methods

EEG signals are composed by a set of specific oscillations known as rhythms. Performing a given mental task (such as motor imagery) makes the amplitude of these different rhythms vary. Moreover, signals such as steady state evoked potentials are defined by oscillations with frequencies synchronized with the stimulus frequency, therefore a lot of BCI systems exploit information embedded in the EEG signal frequency. Two main techniques, which are closely related, are used: band power features and power spectral density features.

3.3.2.1 Band power features

Computing a band power feature consists in band-pass filtering a signal in a given frequency band, then in squaring the filtered signal and finally in averaging the obtained values over a given time window [Pfurtscheller and Neuper, 2001, Boostani et al., 2007]. It is also possible to log-transform this value in order to have features with a distribution close to the normal distribution. Band power features are generally computed for several frequency bands previously determined according to the mental states to be recognized. Such features have been notably used with success for motor imagery classification [Pfurtscheller and Neuper, 2001] but also for the classification of cognitive processing tasks [Palaniappan, 2006, Yamanaka and Yamamoto, 2010].

3.3.2.2 Power spectral density features

Power Spectral Density (PSD) features, sometimes simply called power spectrum, carry information about power distribution of a signal between the different frequencies. PSD features of a signal $x(n)$ can be computed by dividing the signal into a number of partially overlapping segments (epocs). The PSD of a single epoc is computed as in (17).

$$S(w) = \frac{1}{2\pi N} \left| \sum_{n=1}^{N} x_n e^{-jwn} \right|^2 \tag{17}$$

PSD features are probably the most used features for BCI, and have proved to be efficient for recognizing a large number of neurophysiological signals [Hu et al., 2006, Unde and Shriram, 2014].

3.3.3. Time-frequency representations

Considering that neurophysiological signals used in a BCI have generally specific properties in both the temporal and frequency domain, other methods, which can be seen as hybrid, have also been used in BCI design. These

methods are based on various time-frequency representations such as the short-time Fourier transform or wavelets, and extract signal information that is both frequential and temporal. The main advantage of these time-frequency representations is that they can catch sudden temporal variations of the signals, while retaining frequential information. On the contrary, pure frequential methods assume the signal to be in a stationary state.

3.3.3.1 Short-time Fourier transform.

Short-Time Fourier Transform (STFT) is used to determine the sinusoidal frequency and phase content of local sections of a signal as it changes over time. It is calculated by multiplying the input signal by a given windowing function w which is non-zero only over a short time period, and then computing the Fourier transform of this windowed signal. In discrete time, the STFT of the signal $x(n)$ is expressed as in (18):

$$\hat{x}(n, w) = \sum_{n=-\infty}^{\infty} x(n)w(n)e^{-jwn}, \tag{18}$$

in this case, m is discrete.

The Time-Frequency representation is obtained by computing this Fourier transform along a sliding window, i.e., for different segments with a given level of overlapping. This method has been successfully used in several BCI studies [Congedo, 2006, Zabidi et al., 2012]. Its main drawback is the use of an analysis window with a fixed size, which leads to a similar frequential and temporal resolution in all frequency bands. It would be beneficial to have a high temporal resolution for high frequencies which describe a fine scale. Wavelet analysis aims at overcoming this drawback.

3.3.3.2 Wavelet transform

Similarly to Fourier transform, wavelet transform decomposes a signal onto a basis of functions [Samar et al., 1999]. This basis of functions is a set of wavelets $\Phi_{a,b}$, each one being a scaled and translated version of the same wavelet Φ known as the mother wavelet, described in (19).

$$\Phi_{a,b}(t) = \frac{1}{\sqrt{a}}\Phi\left(\frac{t-b}{a}\right). \tag{19}$$

The wavelet transform can then be written as (20):

$$W_x(s, u) = \int_{-\infty}^{\infty} x(t)\Phi_{u,s}(t)dt, \tag{20}$$

where s and u are respectively the scaling and translating factor. The advantage of wavelets is that they make it possible to analyze the signal at different scales

simultaneously. The resolution depends on the scale. As such, high frequencies, which correspond to a fine scale, can be analyzed with a high temporal resolution whereby low frequencies can be analyzed with a high frequential resolution. These factors make wavelets a very useful tool for analyzing EEG signals [Samar et al., 1999]. Various kinds of wavelets have been used for BCI, such as Daubechies wavelets [Michahial et al., 2012, Shete and Shriram, 2014], Morlet wavelets [Bajaj and Pachori, 2014, Qassim et al., 2012] or Mexican hat wavelets [Bajaj and Pachori, 2014, Rafiee et al., 2009].

3.4. Feature selection

BCI feature vectors are often of very high dimensionality [Suk and Lee, 2010]. Several features are generally extracted from several EEG channels and from several time segments, before being used as a single feature vector. Moreover, the training sets, i.e., the example data for each class, are generally small, as the training process is time consuming and relatively uncomfortable for subjects. BCI systems are commonly affected by a well known problem – the "curse-of-dimensionality" [Verikas and Gelžinis, 2008]. This problem emerges, because the amount of data required to describe the different BCI classes increases exponentially with the dimensionality of the feature vector. If the number of training data is small relatively to the number of features, the classification algorithm will give bad results. It is recommended to use at least 5 to 10 times more training data per class than the number of features [Raudys and Jain, 1991]. To tackle this problem, it is often necessary to use dimensionality reduction methods such as principal component analysis [Talukdar et al., 2014]. For a comparison of feature selection methods, see [Lan et al., 2010]. These methods make it possible to work with a set of features with a much smaller size than the original set which generally leads to better performances.

3.5. EEG signal analysis procedure

An individual's brain wave patterns are unique. The brain's activity is mostly generated by internal mechanizms, not external stimulus.

Due to its ability to perform spontaneous activity, the brain does not simply process, but also generates information. Systems with such features are often called complex. This term does not simply mean complicated but implies a nonlinear relationship between constituent components, history dependence, fuzzy boundaries, and the presence of amplifying–damping feedback loops. Complexity can be formally defined as nonlinearity.

Oftentimes, not only does complexity characterize the system as a whole, but also its parts (e.g., neurons) are complex systems themselves, forming hierarchies at multiple levels. All these features are present in the brain's dynamics because the brain can also be considered to be a complex system [Buzsaki, 2006].

The EEG signal may also be considered as chaotic [Klonowski et al., 1999, Ulbikas et al., 1998], the nonlinear dynamics and chaos theory methods can be applied for analysis and classification of the EEG data. To apply chaos theory in EEG analysis it is necessary to reconstruct attractors from EEG signals.

Nonlinearity of the EEG and other types of biosignals has been studied by applying nonlinear analysis methods such as Detrended Fluctuation Analysis (DFA) [Krishnam et al., 2005], Poincare Plots [Chatlapalli et al., 2004], reconstruction of Local Phase Space, segmentation of the EEG signal into stationary fragments [Ursulean and Lazar, 2007], application of non-linear operators to the EEG time series [Martišius et al., 2012].

An EEG signal consists of many components with different characteristics. A large amount of data received from even one single EEG channel presents a difficulty for interpretation. The EEG recorded brain waves originate from a multitude of different neurons from various regions of the brain. Similar coherent electrical activity can be picked up in nearby electrodes. Neurons produce electrical contributions or components that can differ by a number of characteristics such as topographic location, frequency, amplitude, latency etc. The volumetric effect of the cerebrospinal fluid, skull and scalp result in a smearing of these electrical components that result in the scalp recorded EEG macropotential [Takeda et al., 2008]. Through the use of high density EEG recordings one can hope to perform an inverse solution to undo as best as possible the volumetric smearing effects of the medium through which the neuro-electric signals permeate.

Some of these signal properties can be controlled willingly by the subject willingly, allowing for control signals to be extracted from the EEG data. Yet it is difficult to differentiate between states of user intention, because the EEG signal is non-stationary, nonlinear and interfered different artifacts. To distinguish user intentions, after EEG acquisition, pre-processing (filtering/denoising), feature extraction and dimensionality reduction is often performed, before machine learning algorithms can be applied.

For the purposes of BCI system design, various EEG signal properties exist. They discriminate brain function and hence can be used as an input mechanism to offer control or communication. EEG signal properties for BCI systems can be categorized into one of the following groups:

1. Spontaneous brain activity
2. Event-related potentials (ERPs)
3. Event-related desynchronization (ERD) and event-related synchronization (ERS).

The following chapters explain EEG analysis methods of these groups.

3.5.1. Analysis of spontaneous EEG

When analyzing spontaneous EEG, there are two possibilities:

1. Search for the appearance of specific EEG components: these include specific spontaneous wave patterns, such as i.e. sleep spindle or a complex of different waves, i.e. K-complex , different from the rest of brain activity.

2. The frequency spectrum is used, and different wave frequencies are observed. The spectral information can be extracted with the use of Fast Fourier Transform (FFT). The frequency spectrum is divided up into separate frequency bands, such as Delta, Theta, Alpha, Beta and Gamma. This division is needed because different wavelengths have a different interpreted meaning [Birbaumer et al., 1999]. As shown in Figure 3.3, the spectral intensity of a typical EEG signal diminishes as frequency rises. The SCPs can have larger amplitudes (10 to over 100 µV) compared to other frequencies, but care must be taken to remove large amplitude muscle movement artifacts.

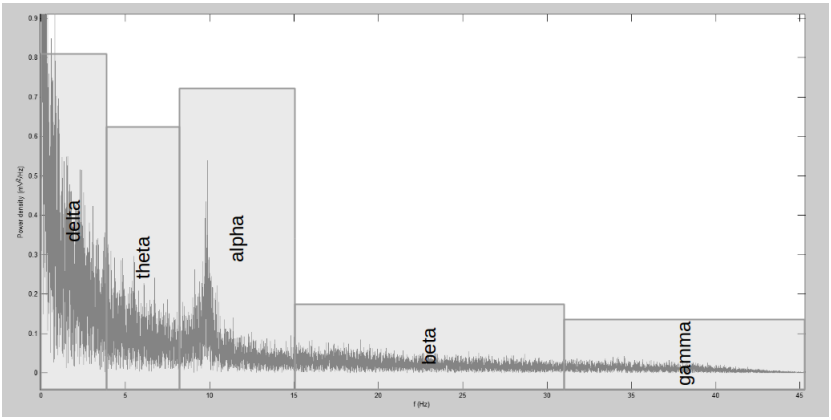


Figure 3.3: A typical frequency spectrum of a healthy subject with eyes open. SCPs normally have larger amplitudes. Local maximums in the spectrum indicate oscillation activity, such as the Alpha rhythm

3.5.2. Event related potential analysis

Since amplitudes of ERP components are often about 1/10 the amplitude of spontaneous EEG components, they are subsequently unrecognizable from unprocessed EEG. A sample of such EEG recording is presented in Figure 3.4. It quickly becomes apparent, that ERP components are indistinguishable from other brain wave activity. It is also clear from the picture, that the signal is not centered around 0, indicating a reference potential drift. Since this drift is different from trial to trial, we can digitally compensate for it, by performing

a detrending step (baseline calibration) on the data, taking the time before stimulus presentation (-150 ms to 0 ms) as a baseline. A baselined signal is presented in Figure 3.5.

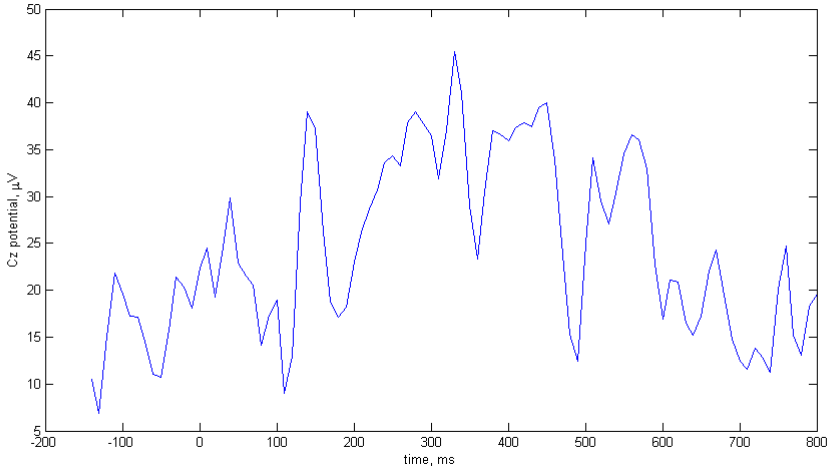


Figure 3.4: A sample of EEG data from a single channel. Time 0 represents stimulus presentation

When trials are detrended, denoising steps are normally performed, in order to remove noisy trials, such as trials where the subject has blinked. Since blinking has a high amplitude, compared to EEG data, this can be accomplished by setting a threshold value on the channels, located above the eyes (Fp1 and Fp2), and removing trials, where the threshold value has been reached.

Once this denoising is performed, the ERP components can be elicited by averaging EEG epochs time-locked to repeated stimulus. The motivation for averaging is that ERPs are time-locked to the stimulus a definite latent period determined from the stimulation time, and they have a similar pattern of response which is predictable under fixed conditions, also the assumption is that the event-related activity, or signal of interest, has a fixed time delay to the stimulus, while the spontaneous background EEG fluctuations are random relative to the time when the stimulus occurred. Averaging across the time-locked epochs highlights the underlying ERP by averaging out the random background EEG activity (similar to additive white noise), thus improving the signal-to-noise ratio [Başar-Eroglu et al., 1996]. These electrical signals reflect only the activity which is consistently associated with the stimulus processing in a time-locked manner. The ERP thus reflects, with high temporal resolution, the patterns of neuronal activity evoked by a stimulus. To get an acceptable signal to noise ratio of the ERP, several (here N) EEG time windows $x_n(t)$ of the same event (of n) are calculated to the so-called grand average $G(t)$, as

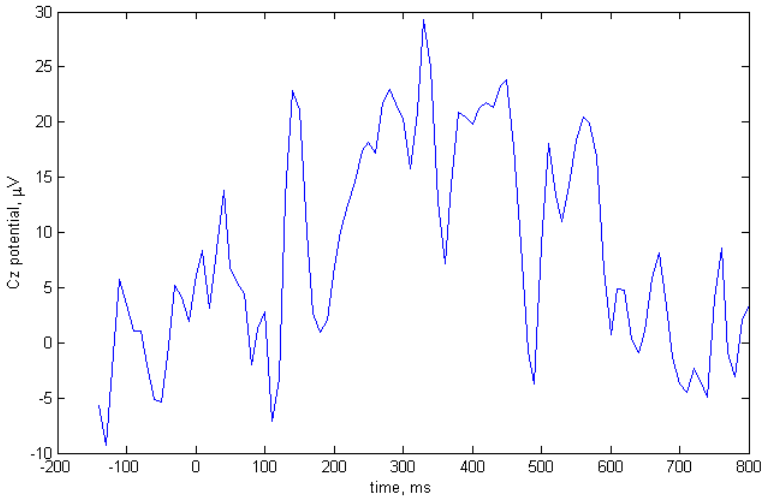


Figure 3.5: A sample of detrended EEG data from a single channel. Time 0 represents stimulus presentation

shown in 21

$$G(t) = \frac{1}{N} \sum_{n=1}^{N} x_n(t), \tag{21}$$

where the signal/noise ratio (SNR) is raised by $\sqrt{N}$ (statistically independent noise process). A grand average of the same channel over 100 trials is presented in Figure 3.6. By comparing the average of trials where the stimulus was presented to the user (called "target" trials), to trials, where no stimulus was presented, one can distinguish the P300 response wave in target trials - the event-specific ERP response potential is visible.

These peaks are composed of interchanging positive and negative activity up to about 0.5 s after the stimulus, and a slow wave, which can last for several seconds. ERPs provide a suitable methodology for studying the aspects of cognitive processes of both normal and abnormal nature, such as neurological or psychiatric disorders.

We can then find channels, where the ERP response is highest, in order to use these particular channels for classification, by plotting all channels, according to their position on the scalp, as shown in Figure 3.7.

3.5.3. Event related synchronization/desynchronization analysis

An internally or externally paced event results not only in the generation of an ERP but also in a change in the ongoing EEG in a form of an event-related

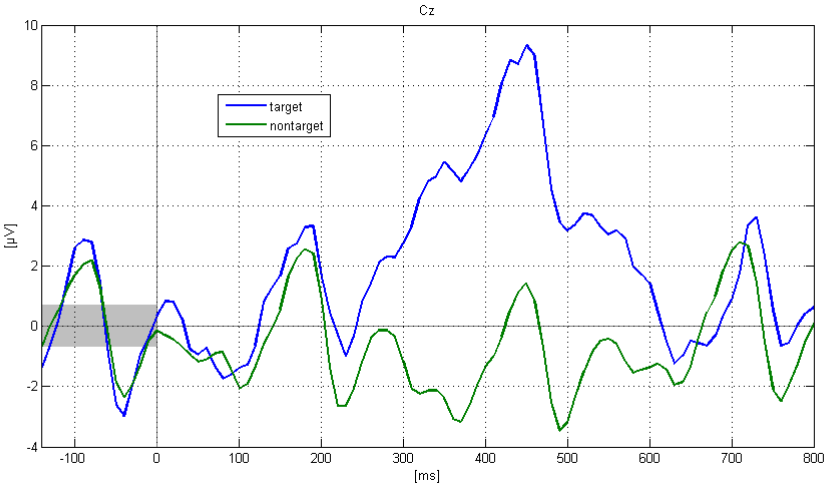


Figure 3.6: A sample of EEG data from a single channel Cz, averaged over 100 trials. Response from times, when stimulus was presented is marked as "target". Response from trials, where no stimulus was presented is marked as "nontarget". The baseline time window, used for detrending, is marked in grey.

desynchronization (ERD) or event-related synchronization (ERS). The ERP on the one side and the ERD/ERS on the other side are different responses of neuronal structures in the brain.

An execution of a task is followed by activity in different brain regions, responsible for the task, leading to uneven activity in the brain - an event called a desynchronization. A desynchronization can be caused by a visual stimulus, attention, planing of movement, imagination and other similar task [Pfurtscheller and Aranibar, 1977]. The EEG channel can be controlled willingly in this way, allowing the person control of a BCI.

ERD is amplitude attenuation and ERS is amplitude enhancement of a certain EEG rhythm. In order to measure ERD or ERS, the power of a chosen frequency band is calculated before and after the event over a number of trials. The average power across a number of trials is then measured in percentage relative to the power of the reference interval. The averaging procedure enhances signal-to-noise ratio. However, this model is an approximation. Some ERPs can be time-locked but not phase-locked to the signal, and therefore cannot be detected by frequency analysis. This means that these ERPs represent frequency specific changes of the continuous EEG activity and may consist of the decrease or increase of power in the frequency bands. This may be considered to be due to a decrease or an increase in synchrony of the underlying neuronal populations. The former is called event-related desynchronization or ERD [Pfurtscheller and Aranibar, 1977], and the latter event-related synchro-

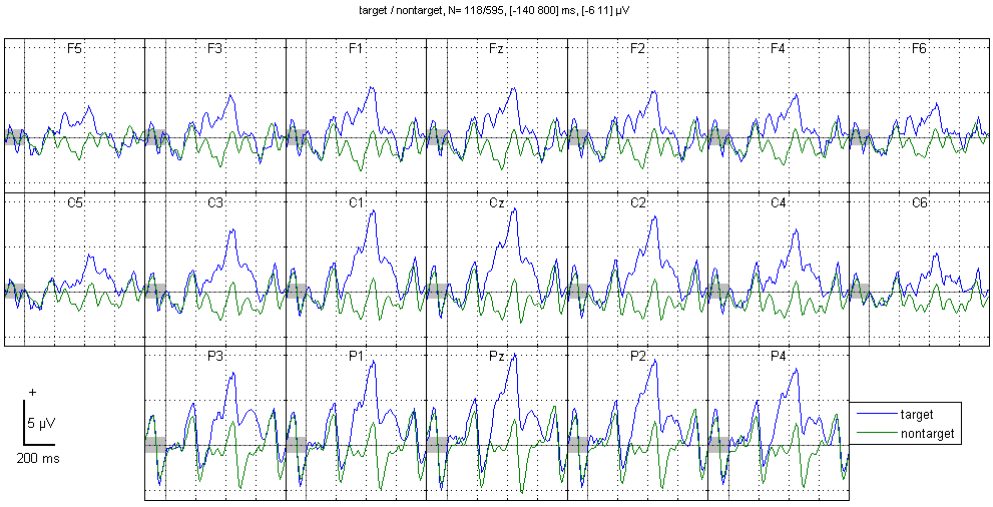


Figure 3.7: ERP responses, averaged over 100 trials, in all EEG channels, displayed by their position on the scalp

nization (ERS) [Pfurtscheller, 1992].

ERD/ERS measurements selected over specific frequency ranges are typically used to produce a spatio-temporal map to visualize the functional behavior of the brain.

One of the basic features of ERD/ERS measurements is that the EEG/MEG power within identified frequency bands is displayed relative (as percentage) to the power of the same EEG derivations recorded during the reference or baseline period before the occurrence of the event.

The method of ERD/ERS analysis is as follows:

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.

Here we provide an example signal processing for imaginative left and right hand movement ERD detection. Motor imagery is a mental process by which an individual rehearses or simulates a given action i.e. left and right hand movement. The idea behind motor imagery is that the motor cortex area of the brain has a synchronized brain pattern in the Beta band. If a person performs right hand movement, the left motor cortex becomes desynchronized. This also holds true, if a person does not perform real physical movement, but imagines performing the action. By left hand motor imagery, the right side of the motor

cortex becomes desynchronized. The data used for this study correspond to the EEG data set IIIb of the BCI competition 2005. The method to compute the time course of ERD is to first compute the spectrum of the signal. The spectrum of a typical trial is presented in Figure 3.8.

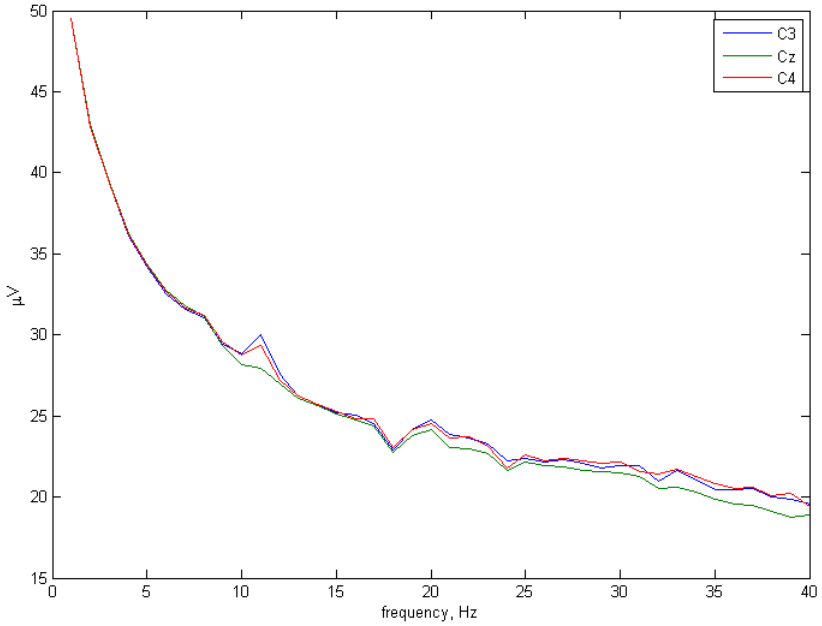


Figure 3.8: Frequency spectrum of a motor imagery trial in electrodes over the motor cortex area (Cz, C3, C4)

By performing noise reduction, the amplitude samples are squared to obtain power sample values. Figure 3.9. Shows all EEG channel plots, according to their scalp location. The power spectrum is presented below each channel.

By examining the band power in channels over the motor cortex, the ERD is clearly visible. Band power can then be used as a feature for classification.

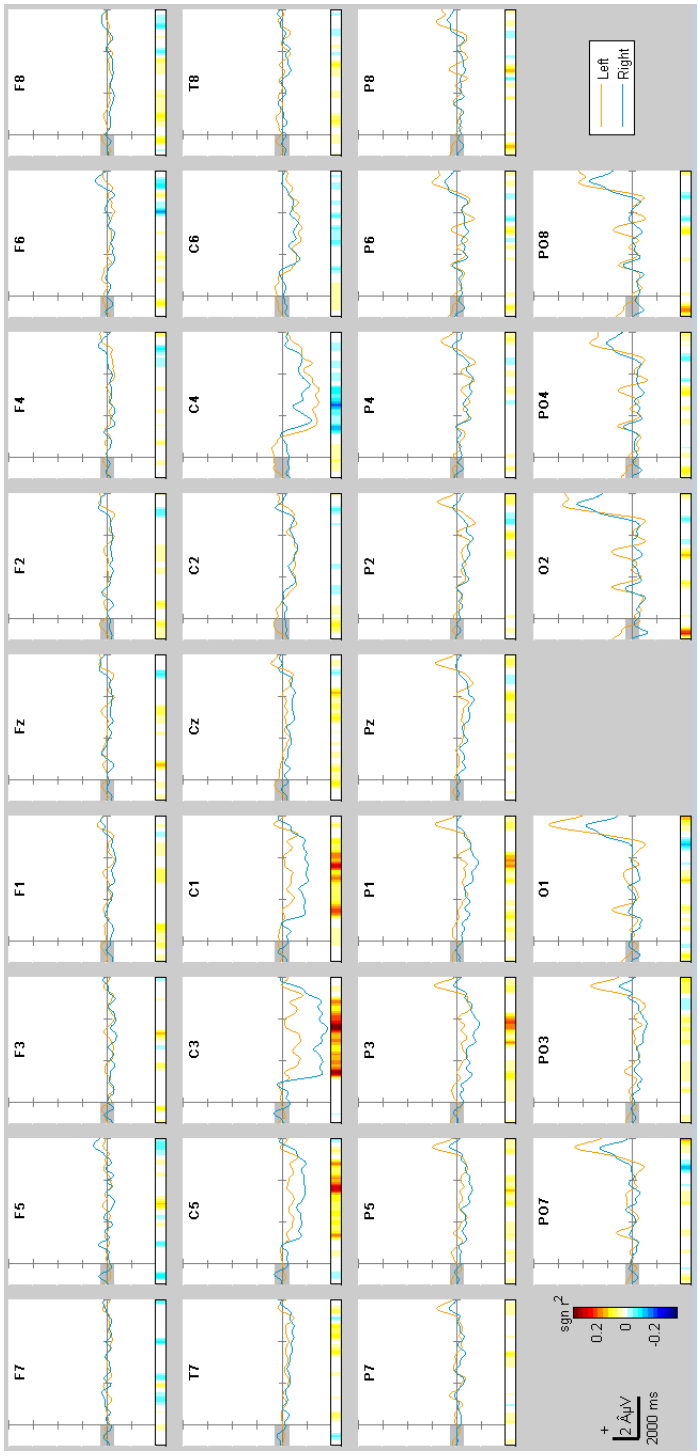


Figure 3.9: Averaged signal of left and right hand movement trials with band power for each EEG channel

3.6. Classification

The key step for identifying neurophysiological signals in a BCI is translating the features into commands [McFarland et al., 2006, Sun et al., 2009]. In order to achieve this step, one can use either regression algorithms, or classification algorithms [Tanaka et al., 2005, Duda et al., 2001, Verikas and Gelžinis, 2008], the classification algorithms being by far the most used in the BCI community [Bashashati et al., 2007, Lotte et al., 2007]. Therefore we focus only on the classification algorithms.

The goal of classification is to assign a correct class label to a previously extracted feature vector. This class represents an intention of the BCI user.

Classification or pattern recognition belongs to a wide field of research known as artificial intelligence (AI). The field of AI focuses on the simulation of human (biological) perspective or cognitive process in machines. The task of classification belongs to a class of AI algorithms used in the simulation of the biological learning process via the use of computer algorithms; the so-called machine learning. Machine learning deals with systems that learn from incoming data, rather than follow programmed instructions.

Algorithms that classify the data are known as "classifiers". Classifiers are able to learn the identification of the feature vector class, thanks to training sets. These sets are composed of feature vectors labeled with their corresponding class. In this section, we first present a taxonomy of the different classification algorithms, and then the main classifier families that are used in the BCI field. These classifiers can be divided into five main categories which are:

- linear classifiers,
- neural networks,
- non linear Bayesian classifiers,
- nearest neighbor classifiers,
- classifier combinations.

3.6.1. Linear classifiers

Linear classifiers are discriminant algorithms that use linear functions to distinguish classes. They are probably the most popular algorithms for BCI applications. Two main kinds of linear classifiers have been used for BCI design, namely, Linear Discriminant Analysis (LDA) [Fisher, 1936] and Support Vector Machines (SVM) [Vapnik, 1998].

3.6.1.1 Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) is a classification method originally developed in 1936 by R. A. Fisher [Fisher, 1936]. It is simple, mathematically robust and often produces models whose accuracy is as good as more complex methods. LDA uses hyperplanes to separate the data representing the different classes. For a two-class problem, the class of a feature vector depends on which side of the hyperplane the vector is (see Figure 3.10).

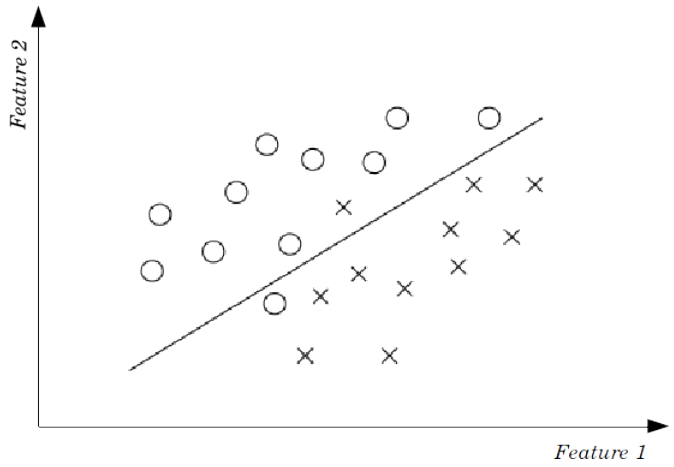


Figure 3.10: A hypothetical classification example with two features. Two data classes, separated by a LDA hyperplane

The separating hyperplane is obtained by seeking the projection that maximizes the distance between the means of classes. To solve an N-class problem, where $N > 2$, several hyperplanes are used. The strategy generally used for multiclass BCI is the "One Versus All" strategy which consists in separating each class from all the others. Such classifier is produces less computational load, is simple to use and generally provides good results [Lotte et al., 2007]. LDA has been used with good results in many BCI systems [Pfurtscheller et al., 1998, Blankertz et al., 2009].

The problem with a standard LDA approach is that it frequently behaves poorly while classifying high dimensionality data. This can be overcome by producing estimates, that are more stable in a high dimensional space. A shrinkage estimate is used for the covariance matrix in the LDA algorithm. This method is called Shrinkage LDA (sLDA). Shrinkage seeks for an estimate of the covariance matrix, such that the expected mean squared error (EMSE) is minimized. A study by [Peck and Van Ness, 1982] indicates that sLDA outperforms the linear algorithm in high dimensions, which is the case in EEG.

A Regularized Fisher's LDA (RFLDA) has also been used

[Müller et al., 2004]. This classifier introduces a regularization parameter C that can allow or penalize classification errors on the training set. The resulting classifier can accommodate outliers and obtain better generalization capabilities. As outliers are common in EEG data, this regularized version of LDA may give better results for BCI than the non-regularized version [Müller et al., 2004].

LDA assumes a normal distribution of the data, with equal covariance matrices for both classes. Although this is not true in many applications, LDA has proven very powerful. One reason is that a linear model is rather robust against noise and most likely will not overfit.

3.6.1.2 Support Vector Machine

Support Vector Machines (SVM) are one of the most popular tools for supervised classification of data. SVM is a binary classification algorithms based on structural risk minimization.

The idea behind an SVM is to map the data into a high dimensional feature space via a nonlinear mapping and to do linear regression in that space. A hyperplane (decision surface) is constructed in this feature space that bisects the two categories and maximizes the margin of separation between itself and those points lying nearest to it (the support vectors) [Muller et al., 2001] (see Figure 3.11). Maximizing the margins is known to increase generalization capabilites [Vapnik, 1998].

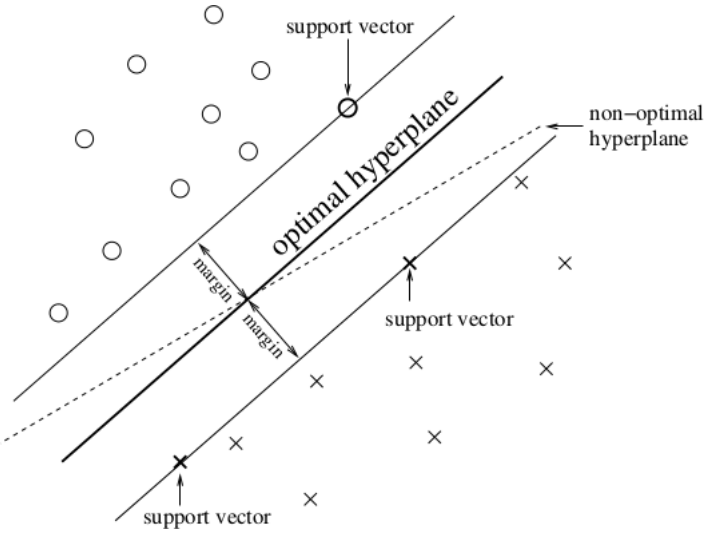


Figure 3.11: Two data classes, separated by an optimal hyperplane, obtained by SVM

This decision surface can then be used as a basis for classifying vec-

tors of unknown classification. Consider an input space X with input vectors $x_i \in X$, a target space $Y = 1, -1$ with $y_i \in Y$ and a training set $T = \{(x_1, y_1), \cdots, (x_n, y_n)\}$. In SVM classification, separation of two classes is done by the maximum margin hyperplane, i.e., the hyperplane that maximizes the distance to the closest data points and guarantees the best generalization on new examples. In order to classify a new point x_j, one has to determine the classification function $g(x_j)$, as in (22):

$$g(x_j) = sgn\left(\sum_{x_i \in SV} \alpha_i y_i K(x_i, x_j) + b\right), \tag{22}$$

where SV are support vectors, $K(x_i, x_j)$ a kernel function, α_i- weights, and b - the bias parameter.

If $g(x_j) = +1$, x_j belongs to class 1, if $g(x_j) = -1$, x_j belongs to class -1, if $g(x_j) = 0$, x_j lies on the decision boundary and can not be classified.

An SVM uses a regularization parameter C which enables accommodation to outliers and allows errors on the training set. Such SVM enables classification using linear decision boundaries, and is known as linear SVM. This classifier has been in a large number of synchronous BCI systems [Wang et al., 2011, Hortal et al., 2013, Martišius et al., 2012, Martišius and Damaševičius, 2012].

It is possible to create nonlinear decision boundaries, with only a low increase of the classifier's complexity, by using the "kernel trick". It consists in implicitly mapping the data to another space, generally of much higher dimensionality, using a kernel function (23). The kernel generally used in BCI research is the Gaussian or Radial Basis Function (RBF) kernel.

$$K(x, y) = \exp\left(\frac{-||x - y||^2}{2\sigma^2}\right), \tag{23}$$

where SV are the support vectors, $K(x_i, x_j)$ is the kernel function, α_i are weights, and b is the offset parameter (bias).

These methods have also given very good results in BCI applications [Bhattacharyya et al., 2011, Hamedi et al., 2014, Wang et al., 2012]. As LDA, SVM have been applied to multiclass BCI problems by combining together multiple two-class SVM [Wang et al., 2012].

SVM have several advantages. Thanks to the margin maximization and the regularization term, SVMs are known to have good generalization properties, and, if trained properly, are insensitive to overtraining and to the curse-of-dimensionality. The quality of the training, however, depends on the given training data, the mapping of data into the feature space, on the kernel function selection for the problem, which must conform with the learning target in order to obtain meaningful results and on the kernel parameters (if any)

[Damasevicius, 2008]. The SVM also has a high computational cost associated with it.

3.6.2. Artificial Neural Networks

The Artificial Neural Network (ANN) is comprised of interconnecting artificial neurons. These neurons, sometimes called nodes, perform simple calculations. The neuron receives input from other neurons or from external sources. Each neuron has an associated weight, which is modified to simulate learning. A single neuron is a relatively simple computational element, so a network of these neurons is most often used for complex tasks, like producing nonlinear decision boundaries [Verikas and Gelžinis, 2008]. The general form an artificial neuron can be described in two stages, as shown in Figure 3.12. In the first stage, the linear combination of inputs is calculated. Each value of input array is associated with its weight value, which is normally between 0 and 1. The summation function will be then performed as in (24)

$$Output = \sum_{i=1}^{n} x_i W_i. \tag{24}$$

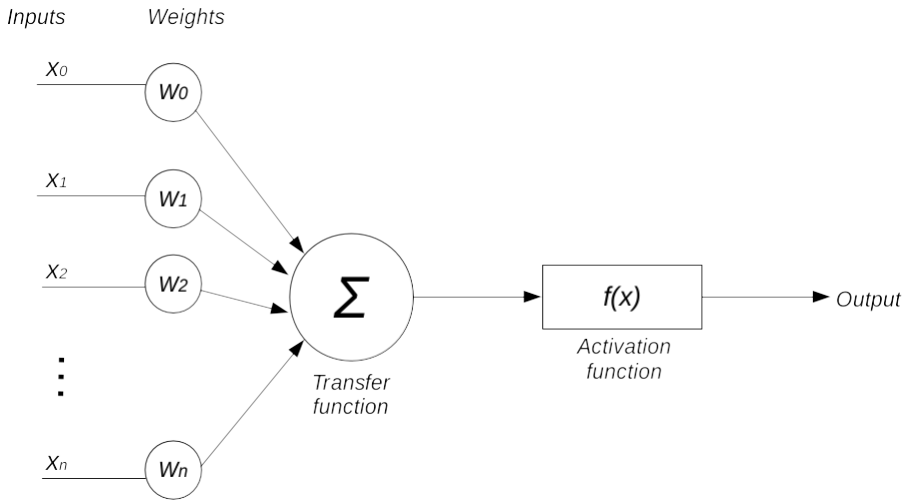


Figure 3.12: A model of an artificial neuron

Two or more neurons can be combined to form a layer. Each input of a layer is connected to inputs of the next layer through neuron weights. Most ANNs have an input layer, an internal, or hidden layer and an output layer.

The most widely used ANNs for BCI is the Multi Layer Perceptron (MLP). It is composed of several layers of artificial neurons: an input layer, one or

several hidden layers, and an output layer. Each neuron's input is connected with the output of the previous layer's neurons whereas the neurons of the output layer determine the class of the input feature vector. This is shown in Figure 3.13.

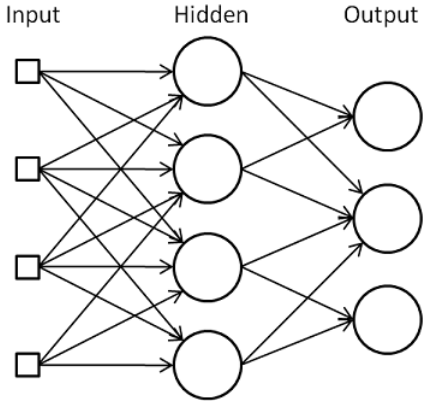


Figure 3.13: The structure of a Multi Layer Perceptron ANN, comprised of an input layer, one hidden layer and an output layer

Multi Layer perceptrons are used in a number of scientific fields, because they are universal approximators: a two-layer MLP can approximate any continuous function. Unlike LDA or SVM, the MLP can classify multiple classes. This makes ANNs very flexible classifiers that can adapt to a great variety of problems. Consequently, MLPs which are the most popular ANN used in classification, have been applied to almost all BCI problems such as binary [Šidlauskas and Martišius, 2013] or multiclass [Yaacob et al., 2013, Hamedi et al., 2014], synchronous [Fu et al., 2011] or asynchronous [Erfanian and Gerivany, 2001, Souza et al., 2012] BCI. However, these classifiers are sensitive to overtraining, especially with noisy, nonstationary data such as EEG. Therefore, careful architecture selection and regularization is required.

3.6.3. Machine learning paradigms

The machine learning process in the BCI domain can be differentiated into two main groups of learning - supervised learning and unsupervised learning.

In a supervised learning BCI system, the classification algorithm is trained to classify mental states by presenting it with a set of training data. The training data is collected in a calibration session. Training data consists of a number of training examples. Each of these examples is a vector of features, extracted from EEG, and a class label that corresponds to desired output value. The algorithm analyzes the training data and produces an inferred function,

which is used to map new examples of data.

There are two basic approaches in choosing the classification function: empirical risk minimization and structural risk minimization [Vapnik, 1998]. Empirical risk minimization seeks the function that best fits the training data. Structural risk minimize includes a penalty function that controls the bias/variance tradeoff. Structural risk minimization seeks to prevent overfitting by incorporating a regularization penalty into the optimization.

If the training algorithm has good generalization, it is able to correctly determine the class label of unseen EEG feature samples. This has been the conventional and most used method of classifier training in BCI. This calibration, however, implies inconvenience to the end users.

Recently, the unsupervised learning approach, has been successfully used in BCI systems [Lu et al., 2009, Vidaurre et al., 2008]. Accurate and robust unsupervised adaptation is considered a key challenge for BCI deployment outside the lab where the intent of the subject and class information is unknown [Krusienski et al., 2011].

Unsupervised learning is trying to find hidden structure in unlabeled data. There is no need for a calibration session. Since the example vector given to the learner are unlabeled, there is no error or reward signal to evaluate a potential solution. The goal in such unsupervised learning problems may be to discover groups of similar examples within the data, where it is called clustering. This distinguishes unsupervised learning from supervised learning. The reader is referred to [Verikas and Gelžinis, 2008] for a detailed analysis and the differences of supervised and unsupervised learning.

3.6.4. Other classification algorithms

Other types of ANN architectures are used in the field of BCI. They deserve mentioning, because some of these algorithms have been specifically created for the purposes of BCI: the Gaussian classifier [Wang et al., 2009a]. Each unit of this ANN is a Gaussian discriminant function representing a class prototype. According to its authors, this ANN outperforms MLP on BCI data and can perform efficient rejection of uncertain samples.

Certain classes of Recurrent Neural Networks, such as Elman's simple recurrent network, are capable of classifying complex, nonlinear, spatiotemporal patterns in EEG. This model can also be allowed to become an autonomous and dynamical system by operating over its own predictions as inputs. When using a network with several hundred neurons these are known to produce good EEG classification accuracy [Møller, 1993, Williams and Peng, 1990]. Other neural network architectures, Bayes classifiers and nearest neighbor classifiers have also been used with good results:

- Voted Perceptron Neural Network [Martišius et al., 2013b,

Šidlauskas and Martišius, 2013],

- Hidden Markov Models [McCormick et al., 2010, Li et al., 2005],
- k-Nearest Neighbor classifiers [Sulaiman et al., 2011, Bahari and Janghorbani, 2013],
- Learning Vector Quantization [Kashtiban et al., 2011],
- Fuzzy ARTMAP Neural Network [Palaniappan and Raveendran, 2002, Palaniappan et al., 2000a, Palaniappan et al., 2000b],
- Time-Delay Neural Network [Rao and Derakhshani, 2005, Barreto et al., 1996],
- Gamma Dynamic Neural Network [Barreto et al., 1996].

3.7. Classifier performance metrics

The most important metrics in BCI system design reported in the literature are accuracy and speed. These are however interdependent and result in inevitable trade-offs, therefore other more complex metrics are often used for classifier evaluation.

Performance of a classifier is it's ability to correctly predict or separate the different classes. The most popular measures, described in short detail below, are the precision, recall, F-measure as well as the area under the receiver operating characteristic curve (also known as Area Under the Curve – AUC) and the Cohen's Kappa.

In a binary classification problem with the classes labeled 'positive' and 'negative', there are several possible classification outcomes: a classifier got a positive class example and labeled it as positive, or it made a mistake and labeled it as negative. Same is true for the negative class. We can then define the following metrics:

- True positives - Tp. The number of positive examples, labeled as positive.
- False positives - Fp. The number of negative examples, labeled as positive.
- True negatives - Tn. The number of negative examples, labeled as negative.
- False negatives - Fn. The number of positive examples, labeled as negative.

Using these metrics, several performance measures can be derived: Precision/Recall, False Positive Rate (FPR)/True Positive Rate (TPR) and F-measure:

Precision or positive predictive value (p) can be defined as the proportion of the true positives against all positive results (both true positives and false positives). It can be calculated according to (25).

$$p = \frac{Tp}{Tp + Fp}. \tag{25}$$

Recall r, also known as TPR the fraction of relevant instances that are retrieved. Both precision and recall are therefore based on an understanding and measure of relevance. It is calculated as in (26)

$$r = TPR = \frac{Tp}{Tp + Fn}. \tag{26}$$

The false positive rate metric can be calculated as in (27) and refers to the probability that a trial gets falsely classified as a user intention.

$$FPR = \frac{Fp}{Fp + Tn}. \tag{27}$$

Both **FPR** and recall represent the rejection performances of the evaluated system as they are independent of the class balance in the data. Precision is linked to the comfort of the BCI system user, as it summarizes how often the BCI system will respond correctly. The values of p and r are tradeoffs: the system with a high r value will generally have low precision and vice versa. To effectively compare different systems with different p and r values, the F-measure is often used.

3.7.1. F-measure

The F-measure (also F_1 score) is a combined measure that assesses the precision and recall values. It considers both the precision p and the recall r of the test to compute the score. The F_1 score can be interpreted as a weighted mean of the p and r, where an $0 \leq F_1 \leq 1$. The balanced F_1 score (a balance between p and r) is the harmonic mean of precision and recall, as in (28).

$$F_1 = 2 \times \frac{pr}{p + r}. \tag{28}$$

The downside of F_1 is that true negative rate (Tn) is not taken into account. It is also biased in some circumstances, therefore, other measures such as the Cohen's kappa, described in section 3.7.4 may be preferable.

3.7.2. Receiver operating characteristic (ROC)

In some studies, receiver operating characteristic (ROC) curves have been used to evaluate the tradeoff between Tp and Fp rates of classification algorithms. A ROC curve is a graphical plot which illustrates the performance of

a binary classifier system as its discrimination threshold is varied. It is created by plotting the fraction of true positives out of the total actual positives, or recall vs. the fraction of false positives out of the total actual negatives (r vs. $1 - p$), at various threshold settings.

The Area Under the Curve (AUC) is of particular interest. The higher the AUC for a given classifier, the better the rejection capabilities of this classifier. The classifier having a higher AUC is considered to perform better. The principal advantage of the AUC is that it is more robust than accuracy in a situation with unbalanced classes. The AUC is a number between 0 and 1. An AUC of 0.5 corresponds to a random classifier. Some of our evaluations in chapters 4.2, 4.5 use AUC as a measure of performance.

3.7.3. Classification accuracy

Classification accuracy is the overall correctness of the classifier and is calculated as the sum of correct classifications divided by the total number of classifications. The accuracy is defined as follows:

$$Acc = \frac{Tp + Tn}{Tp + Tn + Fp + Fn}. \tag{29}$$

The accuracy paradox for predictive analytics states that predictive models with a given level of accuracy may have greater predictive power than models with higher accuracy. Therefore some authors state that the accuracy metric should be used with caution and other metrics such as precision and recall should be favored [Bruckhaus, 2007], especially if the data classes are unbalanced.

There is a need to specify a baseline for the performance of any classification algorithm, because different task might produce very different results i.e. an accuracy of 95 % may be very good in one task, yet extremely bad in another. In a BCI application, Allison et al. places this accuracy baseline in a range of 70%. [Allison and Neuper, 2010]. Another baseline to consider is chance classification accuracy, i.e, in a two class problem, an accuracy of 50% is considered to be a random classification.

3.7.4. Cohen's Kappa statistic

Agreement Statistics argue that accuracy does not take into account the fact that correct classification could be a result of coincidental concordance between the classifier's output and the class label. Cohen's Kappa statistics (κ), defined in (30) corrects this problem [Cohen, 1960]. κ is a measure of agreement normalized for chance agreement. It is a discrete multivariate technique of use

in accuracy assessment.

$$\kappa = \frac{P_A - P_E}{1 - P_E} \tag{30}$$

Where P_A is the probability of overall agreement over the label assignments between the classifier and the true process, P_E represents the probability of chance agreement and $0 \leq \kappa \leq 1$.

- $\kappa = 1$ indicates perfect agreement (a classifier that always classifies correct);
- $\kappa > 0.80$ represent strong agreement and good accuracy;
- $0.40 \leq \kappa \leq 0.80$ is moderate;
- $\kappa < 0.40$ is poor;
- $\kappa = 0$ indicates chance agreement (a random classifier);
- $\kappa = -1$ indicates perfect disagreement.

3.7.5. Cross-validation

With machine learning paradigms it is always better to have more test data in order to obtain and verify a classifier. Ideally, the algorithm would have access to the entire population or a lot of representative data from it. This is usually not the case, and the limited data available has to be reused in order to be able to estimate the error of the classifiers as reliably as possible, i.e., to be reused in order to obtain sufficiently large numbers of samples. The use of re-sampling allows obtaining more stable estimates of an algorithm's performance, and can also enhance replicability of the results.

In k-fold Cross-Validation, the data set is divided into k folds and at each iteration, a different fold is reserved for testing and all the others, used for training the classifiers. It is the best known and most commonly used resampling technique. The testing sets are independent of one another, as required, by many statistical testing methods, but the training sets are highly overlapping. This can affect the bias of the error estimates [Raudys and Jain, 1991].

It is very difficult to qualitatively measure classification performance. Most of the measures focus on different aspects of machine learning, i.e., the precision and recall are quite different, and often good results on one yield bad results on the other. Global statistics, such as accuracy, AUC and F-measure evaluate different aspects of machine learning. All these metrics are used when the general usefulness of a new algorithm is assessed.

For many systems, including BCI, achieving 100 % accuracy is never possible. This is not always the fault of the algorithms: there are algorithms which

can achieve excellent classification performance. This is associated with the concept of ceiling effect, in which an independent variable no longer has an effect on a dependent variable after reaching a certain level, i.e., the ceiling. It is the absolute maximum performance an algorithm will ever be able to achieve.

Knowing this, the performance of any machine learning algorithm will lie between the baseline (chance classification level) and a ceiling (unknown but $<$ 100%).

3.7.6. BCI performance measures

Performance measures are predominantly focused on information transfer capabilities at this premature stage of BCI development. As BCI technology evolves, there will be a much wider range of general performance metrics for BCIs. These may include for example the following:

1. Channel dimensionality,
2. Bit rate of data transfer (channel capacity) – determined by accuracy and speed,
3. Degree of bidirectional control (feedback),
4. Short- and long-term reliability,
5. Cost effectiveness,
6. User-friendliness (comfort, portability, ease-of-use),
7. Training requirements,
8. Staff and service requirements (required assistance).

One of the most important yet subjective performance metrics involves a measure of the quality-of-life improvement for the user. The target population's lack of conventional communication ability makes it initially difficult to quantify this metric but perhaps in the future, this could become the ultimate performance metric.

3.8. Feedback and application

After classification has been performed, the specific classes can be associated with a control command to a given application.

These applications can be divided into two main categories - medical applications, and non-medical applications, such as multimedia or virtual reality. The first category includes:

- **Rehabilitation and prosthetic device control.** BCI technology is primarily designed for patients with moderate to severe movement disabilities. Although rehabilitation is impossible in some diseases, such as ALS, some of the patients, i.e. stroke patients, can sometimes regain some or all lost motor control with effective rehabilitation. Motor imagery BCI can be used as a means for rehabilitation.

  In studies [Pfurtscheller et al., 2005, Wang et al., 2009b, Ang et al., 2012] among others, patients have tried to grasp objects using BCI controlled robotic prosthetic hands. Robotic arms provided feedback for the patients, aiding their rehabilitation. While rehabilitation results show potential, robotic prosthetic limb control requires a number of control commands, not achievable by BCI systems. This is covered in more detail in chapter 2.9. The experiments, therefore, are mostly limited to 1-D or 2-D control.

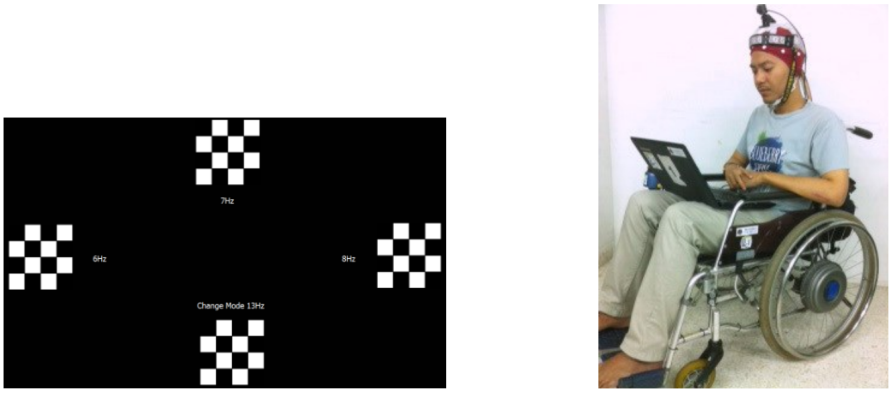
- **Assistive mobility.** The most beneficial devices for disabled people are those that let them regain mobility. This is achieved by providing wheelchair control, by means of BCI. Many BCI paradigms have been exploited, such as SSVEP [Punsawad and Wongsawat, 2013, Muller et al., 2011], ERD/ERS [Huang et al., 2012], MI [Reshmi and Amal, 2013, Carrino et al., 2012] and hybrid methods [Li et al., 2013].

  In a study by Punsawad et al. [Punsawad and Wongsawat, 2013], subjects used SSVEP and motor imagery to control a wheelchair shown in Figure 3.14. Subjects achieved approximately 90 % accuracy in the SSVEP system, and 85.62 % accuracy in motor-imagery tasks.

- **BCI-driven spelling devices** are used to spell letters or words, allowing for disabled communication. The P300 speller is one of the most famous BCI paradigms [Farwell and Donchin, 1988]. In this setup, the subject is seated in front of a screen on which a 6×6 matrix is displayed. The feedback interface is presented in Figure 3.15.

  In this application, a row or a column of the matrix is highlighted every 125 ms. The user is asked to look at and draw his attention to the letter he wants to select, and to count the number of times the desired letter is highlighted. The highlighting of the desired letter triggers a P300 in the user's EEG. Detecting the presence of the P300 makes it possible to find which are the line and column that contain the desired letter. This speller could enable its users to spell up to 7.8 letters/min. [Farwell and Donchin, 1988].

  Another recent feedback paradigm is the Hex-o-spell [Blankertz et al., 2006b] speller, displayed in Figure 3.16. Motor



(a) The SSVEP stimulator **(b)** The BCI controlled wheelchair

Figure 3.14: Subject, controlling a robotic wheelchair by SSVEP. By focusing attention on a blinking stimulus in (a), the subject can issue forward/left/right/reverse commands for the wheelchair (b). [Punsawad and Wongsawat, 2013].

imagery is used for control.

In this setup, the letters are displayed in 6 groups, consisting of 5 letters each. For the selection of a symbol there is an arrow in the center of the circle. By imagining a right hand movement the arrow turns clockwise. An imagined foot movement stops the rotation and the arrow starts extending. Once it touches the hexagon, the selection is performed. Then the letters of that hexagon are moved into individual hexagons and another selection step is performed to select the letter or symbol. The typing speed is subject-depended but Blankertz et al. achieved 7.6 letters/min. [Blankertz et al., 2006b].

- **BCI controlled web and music browsers.** Internet access has become the main source of communication on a global scale. BCI technology enables the development, to make the internet accessible for the disabled. As more aspects of daily life become accessible online (education, retail,personal finance, or business), the potential benefit of connectivity also increases. In a study by E.M. Mugler et al. [Mugler et al., 2010] patients used the P300 paradigm to navigate text, browse forward and backward, use bookmarks and spelling text. The feedback interface is presented in Figure 3.17. The study used a 16 channel g.USBamp amplifier from g.tec. Reported accuracy average is 77.4 %. Another study [Pinegger et al., 2015], using the g.Nautilus amplifier with 32 electrodes reports reaching accuracies between 84.9 % and 96.5 % in performing similar internet tasks with automatic link and text detection. In a study

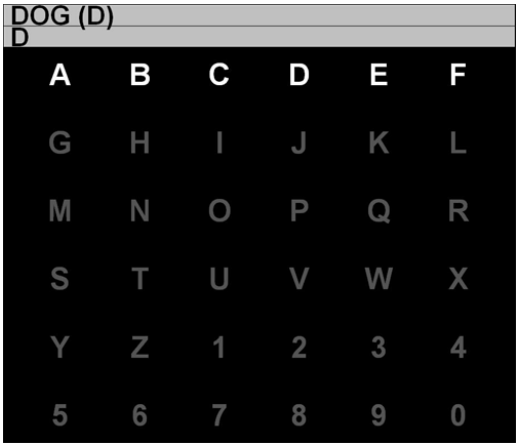


Figure 3.15: The standard P300 feedback interface

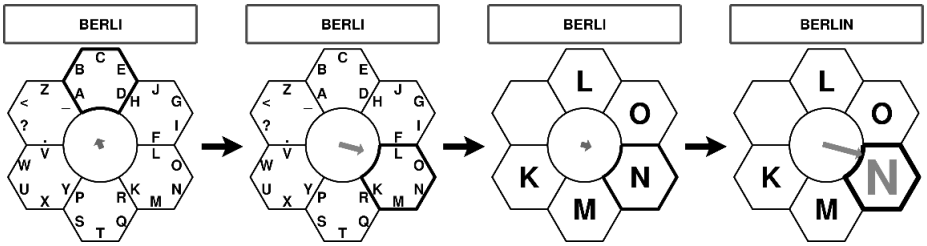


Figure 3.16: The mental typewriter 'Hex-o-Spell'. The two states classified by the BBCI system control the turning and growing of the gray arrow respectively (see also text). Letters can thus be chosen in a two step procedure [Blankertz et al., 2006b].

[Teo et al., 2006], the P300 BCI is used as a music selection and media interface with promising results.

- **Mental state recognition.** Work in this area deals with the recognition of mental states, such as attention levels, to treat attention deficit disorder patients [Rohani et al., 2014], workload and fatigue [Roy et al., 2013], useful for an operators cognitive state assessment.

  Rohani et al. [Rohani et al., 2014] used the P300 paradigm to create a game for attention training. The game is intended for subjects, suffering from attention deficit disorder. Accuracy results are 70.66 %.

  Roy et al. [Roy et al., 2013] measured mental fatigue by exploiting adaptive EEG band power features. The classification accuracy levels are reported to be 98.04 % for time-on-task mental fatigue and 65.51 % on workload fatigue recognition.

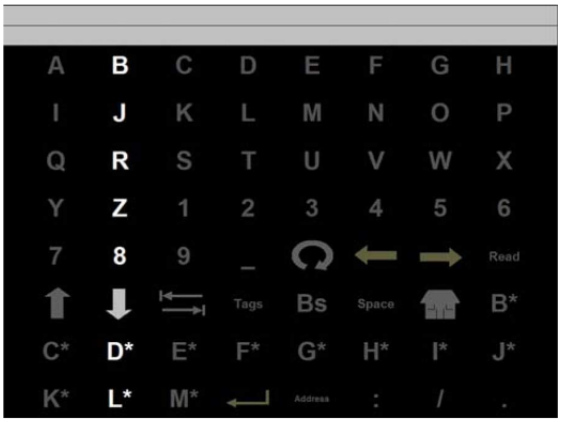


Figure 3.17: Character matrix for P300 Browser [Mugler et al., 2010].

Although BCI technology is mainly designed with disabled people in mind, it can also be beneficial to healthy subjects. EEG is particularly suited for this purpose, because it is non-invasive, portable, has a good temporal resolution of a few milliseconds and is relatively low cost. Therefore non-medical applications include:

- **Gaming**. All BCI paradigms have been exploited for gaming purposes. BCI is used either as a primary means to control the game, or an extra channel for in-game communication, to perform certain user actions, whereas the game is primarily controlled by traditional means.

  Game examples include a 3-class MI-based asteroid-dodging game, described in [Coyle et al., 2011] and a BCI control interface for a popular game "World of Warcraft", [Kapeller et al., 2012]. SSVEP based games include a 2 class game called MindBalance [Lalor et al., 2005a] and a P300 based the MindGame [Finke et al., 2009].

- **Virtual reality**. Most existing works focus on either rotating the virtual camera or traveling in the virtual environment. Pineda et al. used a BCI based on the mu rhythm to interact with a "First Person Shooter" video game [Pineda et al., 2003]. A high mu rhythm level triggered left camera rotation, whereas low mu levels triggered right rotations. Other commands in the game were issued by using the keyboard. No accuracy measurement was performed in this study.

3.9. Conclusions

The third chapter reviews standard techniques for EEG signal processing and feature extraction. These include the use of temporal and spatial filters,

Figure 3.18: The players Avatar on a tightrope between two SSVEP checkerboards. The player must correct the avatar's fall by focusing on the appropriate symbol [Lalor et al., 2005a]

such as CSP and Fourier transform filters. We have also reviewed feature extraction methods, which include temporal filters, Hjorth parameters, frequency domain features and wavelet transform features.

All of these techniques have been shown to be of use in BCI applications and produce good results.

We then review classification algorithms for BCI. These include linear classifiers and artificial neural networks. Simple classifiers are shown to suffer from high dimensionality of features and therefore have to be chosen according to the feature extraction algorithm being used in a particular case.

In this chapter we provide a review of classifier and system performance metrics. We conclude that simple metrics, such as overall accuracy, while being directly linked to system usability, is not always sufficient. This is especially true for unbalanced datasets. Therefore, more complex metrics, such as F-measure and Kohen's Kappa statistics are used.

Finally, we reviewed various BCI applications, both medical and non-medical, used for virtual reality and gaming.

This chapter shows that despite a large number of studies and research done in various fields for designing BCI systems, no distinct single-best method is developed. The BCI community has highlighted the need for more efficient algorithms. This point is addressed in the next chapter.

4. METHODS

In our experiments, we predominantly focus on EEG signal preprocessing and feature extraction algorithms. In this chapter, we propose several novel methods for BCI applications. These methods have been developed and published in [Martišius et al., 2012, Martišius et al., 2013b, Martišius et al., 2013a, Martišius and Damaševičius, 2012, Šidlauskas and Martišius, 2013, Damaševičius et al., 2014, Vasiljevas et al., 2012].

The developed algorithms are tested on standard datasets, to allow for a comparable result. All methods in this chapter are used in off-line EEG analysis.

This chapter is organized as follows: In section 4.1 we provide a description of the standard datasets used. Section 4.2 describes a novel signal denoising scheme based on shrinkage functions. Section 4.3 presents a novel feature extraction algorithm for BCI which uses fractional delay time embedding. Results, obtained by the algorithm can be used as features for a classifier. In section 4.4, we present a feature extraction method based on the Wave Atom transform. This method describes the signal with a few components, allowing for fast classifier training and good classification results. Section 4.5 focuses on the use of nonlinear operators for EEG feature extraction. Various operator performance for BCI feature extraction is compared. We also propose a novel operator, called HMPO, which stands for Homogenous Multivariate Polynomial Operator. In section 4.6 we propose the time-bound training method for ANN networks and show that it can be used for fast classifier retraining, needed in real-world applications. Finally, a discussion of the obtained results and conclusions are provided.

4.1. EEG datasets

There is much talk in the literature for a standardized framework to facilitate the direct performance comparison of BCI systems and their individual components [Thompson et al., 2013]. The development of benchmarks would be extremely valuable, enhancing the BCI research community's ability to objectively compare BCI technologies.

The BCI competition, hosted by the Berlin Brain Computer Interface (BBCI) project, can be considered to be such a benchmark. BCI Competition is an open competition which aims at evaluating various approaches used in BCIs and comparing them on the same data set in order to obtain a reliable measurement of performance for each algorithm. It is an attempt to solve the problem of comparing signal processing methods that are published, but their accuracy is verified on different data sets or they use different performance measures.

Four editions of the BCI Competition were organized (the last one in 2008) and each edition consisted of 3–5 data sets for which a different classification

task was provided. Each data set is freely available on the competition web page http://bbci.de/competition and is divided into a training set and a test set. The goal of the competitors is to provide programs that would be able to correctly classify the data from the test set.

4.1.1. BCI competition II dataset Ia and Ib

The datasets were taken from a healthy subject. The subject was asked to move a cursor up and down on a computer screen, while his cortical potentials were taken. During the recording, the subject received visual feedback of his slow cortical potentials (Cz-Mastoids). Cortical positivity lead to a downward movement of the cursor on the screen. Cortical negativity lead to an upward movement of the cursor. Each trial lasted 6s.

During every trial, the task was visually presented by a highlighted goal at either the top or bottom of the screen to indicate negativity or positivity from second 0.5 until the end of the trial. The visual feedback was presented from second 2 to second 5.5. Only this 3.5 second interval of every trial is provided for training and testing. The sampling rate of 256 Hz and the recording length of 3.5s results in 896 samples per channel for every trial. There is a total of 268 trials.

Data:

- Amplifier: PsyLab EEG8
- A/D-converter: Computer Boards PCIM-DAS1602/16 bit
- Amplitude range: +/-1000 µV
- Sampling rate: 256 S/s
- EEG data was taken from the following positions:
- Channel 1: A1-Cz (10/20 system) (A1 = left mastoid)
- Channel 2: A2-Cz
- Channel 3: 2 cm frontal of C3
- Channel 4: 2 cm parietal of C3
- Channel 5: 2 cm frontal of C4
- Channel 6: 2 cm parietal of C4
- All values in µVolt.

4.1.2. BCI competition II dataset IV

This dataset was recorded from a healthy subject during a no-feedback session. The subject sat in a chair, relaxed arms resting on the table, fingers in the standard typing position at the computer keyboard. The task was to press with the index and little fingers the corresponding keys in a self-chosen order and timing 'self-paced key typing'. The experiment consisted of 3 sessions of 6 minutes each. All sessions were conducted on the same day with some minutes break inbetween. Typing was done at an average speed of 1 key per second.

Data:

- Amplifier: NeuroScan
- 28 EEG channels in 10/20 system
- Sampling rate: 1000 S/s
- EEG data was taken from the following positions: F3, F1, Fz, F2, F4, FC5, FC3, FC1, FCz, FC2, FC4, FC6, C5, C3, C1, Cz, C2, C4, C6, CP5, CP3, CP1, CPz, CP2, CP4, CP6, O1, O2
- Band-pass filter between 0.05 and 200 Hz.
- All values in µVolt.

4.2. Class-Adaptive Signal Denoising

As mentioned in Chapter 3, the EEG signal usualy undergoes a denoising step, before feature extraction and mental task classification can be performed.

Linear methods, such as FIR filters have some side effects while removing noises. Therefore non-linear denoising methods in the wavelet domain have been an active area of study. Wavelet denoising techniques are based on some threshold value for removing small coefficients of detail sub-bands and preserving large coefficients, because small coefficients are usually noisy and large coefficients contain the main signal features. Thresholding needs a function to decide how improve coefficients. Therefore, estimating threshold and determining thresholding rules are still challenging problems in wavelet denoising [Norouzzadeh and Jampour, 2011].

There is a number of shrinkage functions developed, such as the soft shrinkage function, hard, firm, hyper-trim shrinkage functions, multi-parameter best-basis thresholding shrinkage functions, Yasser shrinkage function et al. All these threshold and shrinkage functions promote the application of wavelets in signal denoising.

Denoising results can be improved, if the thresholds at each frequency scale are chosen optimally. Such subband-adaptive denoising is used by modern wavelet-based denoising algorithms such as SureShrink

[Donoho and Johnstone, 1995] and BayesShrink [Chang et al., 2000]. Optimization of the class-to-class distance for two-class discrimination using Fisher distance is described by Aldjem [Aladjem, 1996]. Mu et al. [Mu et al., 2009] also use Fisher distance in feature selection algorithm and apply it for classification of EEG data.

The novelty of this algorithm is the proposed Class-Adaptive denoising method to select optimal parameter value(s) of a standard shrinkage function by maximizing class distance between frequency domain components of the positive and negative data classes.

For our analysis, we classify shrinkage functions depending upon the dimensionality of their parameter space as: single-parameter, two-parameter, three-parameter and multi-parameter shrinkage functions. Dimensionality of the parameter space is important for the selection of an optimization method to find best parameter values. Below we provide a short description and analysis of some of these functions.

4.2.0.1 Single parameter shrinkage function overview.

In 1995 Donoho and Johnston developed the hard (31) and the soft (32) shrinkage functions [Donoho, 1995].

$$\hat{y} = \begin{cases} 0, & \text{if } |y| \leq \lambda, \\ y, & \text{if } |y| > \lambda. \end{cases} \tag{31}$$

$$\hat{y} = \begin{cases} 0, & \text{if } |y| \leq \lambda, \\ y, & \text{if } |y| > \lambda, \\ y + \lambda & \text{if } |y| < -\lambda. \end{cases} \tag{32}$$

Here y is the noisy value, $\hat{y}$ is the shrinked value, and λ is universal threshold. The derivation of standard soft shrinkage function is not continuous. Both hard and soft shrinkages have advantages and disadvantages. The soft shrinkage estimates tend to have a bigger bias, due to the shrinkage of large coefficients. Due to the discontinuities of the shrinkage function, hard shrinkage estimates tend to have a bigger variance. In other words, it will be sensitive to small changes in the signal.

Norouzzadeh and Jampour [Norouzzadeh and Jampour, 2011] propose shrinkage function (33):

$$\hat{y} = y - \frac{5 \times \lambda^2 y}{y^8 + 5 \times \lambda^2}. \tag{33}$$

This function is continues, differentiated and has a higher order derivation. Therefore it is suitable for any gradient decent learning algorithm.

Poornachandra and Kumaravel [Poornachandra and N., 2006] propose a hyper-trim shrinkage function (34).

$$\hat{y} = \begin{cases} 0, & \text{if } |y| \leq \lambda, \\ sgn(y) \times \sqrt{y^2 - \lambda^2}, & \text{if } |y| > \lambda. \end{cases} \tag{34}$$

This function generalizes hard and soft shrinkage proposed by Donoho and Johnston. The denoising method is based on Stein's unbiased risk estimation (SURE) and on a new class of shrinkage function. The proposed new class of shrinkage function has continuous derivative. The main objective of this function is to reduce the minimum mean square error between original signal y and the denoised signal $\hat{y}$. The authors note that this shrinkage function was simulated and tested with ECG signals added with standard Gaussian noise and gave better mean square error performance over conventional wavelet shrinkage methodologies.

4.2.0.2 Two-parameter shrinkage function overview.

Poornachandra and Kumaravel [Poornachandra and Kumaravel, 2008] also propose a hyper shrinkage function (35).

$$\hat{y} = \begin{cases} 0, & \text{if } |y| \leq \lambda_1, \\ tanh(\lambda_2 \times y), & \text{if } |y| > \lambda_1. \end{cases} \tag{35}$$

where λ_1 and λ_2 are the parameters of the functions.

This function is designed to filter power line frequency noise without using the conventional finite impulse response (FIR) filter. Hyper shrinkage is an optimized thresholding scheme based on alpha trim thresholding. The major advantage of hyper shrinkage is its nonlinearity, that is, the function in wavelet domain tends to keep a few larger coefficients representing the function while the noise coefficients tend to be reduced to zero.

Another two-parameter shrinkage function (36) is proposed by Mrazek et al [Mrázek et al., 2005]:

$$\hat{y} = y \left(1 - 2 \times 10^{-\frac{2y^2}{\lambda_1^2}} + 10^{-\frac{2y^2}{\lambda_2^2}} \right) \tag{36}$$

here λ_1 and λ_2 are the parameters of the functions.

This shrinkage functions has been shown by the authors to be competitive with the best previously known shrinkage methods when applied to signal denoising with multiscale wavelet procedures.

4.2.0.3 Three-parameter shrinkage function overview.

Ying and Yusen [Ying and Yusen, 2009] propose a generalization of soft, firm and Yasser shrinkage function, as in (37).

$$\hat{y} = \begin{cases} 0, & \text{if } |y| \leq \lambda_1, \\ sgn(y) \times \left[\frac{|y-\lambda_1|^{\lambda} \times \lambda_2}{|\lambda_2 - \lambda_1|^{\lambda}} \right], & \text{if } \lambda_1 < |y| \leq \lambda_2, \\ y, & \text{if } |y| > \lambda_2. \end{cases} \tag{37}$$

here γ, λ_1 and λ_2 and are the parameters of the functions. This shrinkage function has been shown to produce very good results with a noisy ECG signal. Not only did it get the highest SNR, but also kept the similarity and smoothness of the denoised signal.

Atto et al. [Atto et al., 2008] propose the smooth sigmoid based shrinkage function (38).

This shrinkage function satisfies the following properties:

1. Smoothness: a smooth shrinkage function induces small variability among coefficients with close values.

2. Penalized shrinkage: a strong (resp. a weak) attenuation for small (resp. large) coefficients is required because small (resp. large) coefficients contain less (resp. more) information about the signal.

3. Vanishing attenuation at infinity: to avoid over-smoothing of the shrinked signal value, the attenuation decreases to zero when the amplitude of the coefficient tends to infinity. In the function (38), the authors put $\beta_{\tau,\lambda} = \beta_{0,\tau,\lambda}$

$$\hat{y} = \frac{sgn(y) \times (|y| - t)_+}{1 + e^{-\tau(|y|-\lambda)}}. \tag{38}$$

Here $(t, \tau, \lambda) \in \mathbb{R}$, are the parameters of the function.

This function is smooth and introduces little variability among coefficients with close amplitudes. The shrinkage is very flexible and allows to (approximately) preserve the statistical properties of the signal.

It is a good trade-of between noise reduction, signal feature enhancement and computational load. Denoising has been shown to achieve good PSNR values and an almost artifact-free denoising.

4.2.0.4 Multi-parameter shrinkage function overview.

Poornachandra and Kumaravel [Poornachandra and N., 2006] propose a sub-band dependent adaptive shrinkage function that generalizes the hard and soft shrinkage functions, as in (39).

A native method of denoising is equivalent to low-pass filtering naturally included in any dyadic wavelet framework. That is, simply discard channels of highest resolution and allow signal in the channel confined to lower frequency. The problem associated with this linear denoising approach is unsuitable, as it does not remove the noise present in the low frequency channel as most of the signals of biomedical origin are of lower frequencies. For any shrinkage scheme to be effective, an essential property is that the magnitude of signal components has to be larger than that of existing noise.

The subband-adaptive nonlinear shrinkage works on hyperbolic function, which outperforms the soft shrinkage.

$$\hat{y} = \begin{cases} \rho \left[\frac{1-\lambda_j^{-2\lambda_j y}}{1+\lambda_j^{-2\lambda_j y}} \right], & \text{if } |y| \geq \lambda_j, \\ 0, & \text{if } |y| < \lambda_j. \end{cases} \tag{39}$$

where:

$$\rho = \frac{\delta}{\max |x|}, \tag{40}$$

ρ is the boundary contraction parameter, which depends on boundary attaining parameter $\delta, 10 > \delta > 1$. λ_j are function parameters for each sub-band j.

4.2.0.5 Evaluation.

The soft (32), hard (31) and various variants of firm shrinkage (such as (37)) are commonly used for denoising, so based on other author's results [Poornachandra and N., 2006, Poornachandra and Kumaravel, 2008, Atto et al., 2008] we can provide the following evaluation of the shrinkage functions. The signal denoised using soft shrinkage tends to have a bigger bias due to the shrinkage of large coefficients, while the discontinuities of the hard shrinkage function produce a bigger variance. Firm shrinkage is less sensitive than hard shrinkage to small fluctuations and less biased than soft shrinkage, however it is more computationally expensive. Hard shrinkage is discontinuous and is not differentiable. Soft shrinkage is continuous, but does not have a first order derivation. Sigmoid based shrinkage is smooth (i.e. induces small variability among data with close values), it produces strong attenuation and is imposed for small data and weak attenuation for large data.

4.2.0.6 Signal denoising using shrinkage functions.

Signal denoising by thresholding is based on the observation that a limited number of the DSP transform coefficients in the lower bands are sufficient to reconstruct the original signal. The key steps of signal denoising using DSP transforms are the selection of shrinkage function and its parameter(s). The

goal of the shrinkage function is to remove noise so that separability of positive class and negative class in a binary classification problem is increased.

Assume that the observed data $X(t) = S(t)+N(t)$ contains the true signal $S(t)$ corrupted with additive noise $N(t)$ in time t. Let $T()$ and $T^{-1}()$ be the forward and inverse transform operators. Let $H(Y, \Lambda)$ be the denoising operator with a set of parameters $\Lambda = (\lambda_1, \lambda_2 \cdots, \lambda_k)$. Then the denoising algorithm is defined as follows:

1. Compute the DSP transform for a noisy signal $X(t)$: $Y = T(X)$;
2. Perform frequency shrinkage in the frequency domain: $\hat{Y} = H(Y, \Lambda)$;
3. Compute the inverse DSP transform to obtain a denoised signal $\hat{S}(t)$ as an estimate of $S(t)$: $\hat{S} = T^{-1}(\hat{Y})$.

Steps 1-3 can be generalized into a single equation as in (41).

$$\hat{S} = T^{-1}(H(T(X), \Lambda)). \tag{41}$$

4.2.0.7 Proposed class-adaptive shrinkage method.

The scheme described in above might not work well in case where signal $S(t)$ and noise $N(t)$ have many different components as is the case with the EEG data. Also the selection of the shrinkage function and its parameters is problematic due to a large number of shrinkage functions proposed in the literature and a large variability in signal data. Therefore, some adaptivity must be introduced when selecting shrinkage function and its parameters. A description of the Class-Adaptive (CA) shrinkage method is provided: Let P and Q be the positive and negative classes of data. Let $D(X_P, X_Q)$ be a distance function between datasets X_P and X_Q belonging to P and Q, respectively. We improve the denoising algorithm by optimizing shrinkage function parameters for each frequency component f of X_P and X_Q, while the aim function in (42):

$$\Psi(X_P, X_Q) = \max_{\Lambda} D(H(T(X_P), \Lambda), H(T(X_Q), \Lambda)). \tag{42}$$

here D is the distance metric between classes.

To calculate a distance between data classes, several distance metrics can be used. We use Fisher distance and Hellinger distance. The Fisher distance [Itoh and Shishido, 2008] between two data classes is calculated as in (43):

$$F = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2}. \tag{43}$$

here μ and σ are the mean and the standard deviation of the class they correspond to.

The squared Hellinger distance [Bar-Yossef et al., 2004] between two data classes with normal distributions is calculated as in (44).

$$H^2 = 1 - \sqrt{\frac{2\sigma_1\sigma_2}{\sigma_1^2+\sigma_2^2}} e^{-\frac{(\mu_1-\mu_2)^2}{4\times\sigma_1^2+\sigma_2^2}}. \tag{44}$$

The proposed CA shrinkage algorithm is as follows:

1. Convert the time domain signals to frequency domain signals using the FFT transform.
2. For each frequency f:
   - maximize distance between frequency components of positive class and negative class with respect to a set of shrinkage function parameters Λ;
   - retain Λ for maximal distance as Λ_{max}.
3. Perform shrinkage of the FFT coefficients using Λ_{max}.
4. Convert the shrinked frequency domain signal to the time domain using an inverse FFT transform.

For this experiment, Data set Ia from the BBCI competition datasets was used. It was first randomly partitioned into 5 parts, and a 5-fold cross-validation was used to evaluate classification results.

FFT was performed on each channel and the shrinkage function parameters were optimized to obtain maximal distance between positive and negative classes. The optimization was performed using Nelder-Mead (downhill simplex) optimization method [Nelder and Mead, 1965] (implemented in Perl's Amoeba package). The classification was performed using SVMPerf [Joachims, 2005] implementation of SVM with linear kernel.

To evaluate the precision of classification, we used the F-measure, Area Under Curve (AUC), and Average Precision (Avg. Prec.) metrics. The results of experiments are given in Table 3 (using Fisher distance) and Table 4 (using Helliger distance).

The achieved results are in line with the best results achieved by other authors while using the BCI competition Ia dataset (e.g., Mench et al. [Mensh et al., 2004] report 88.7 % correct classification rate using Thomson multitaper method).

Table 3: Experimental results using Fisher distance

CA shrinkage function	No. of Parameters	F–measure	AUC	Avg. Prec.
Not applied	-	0.80	90.06	91.45
Hard	1	0.84	91.53	92.60
Soft	1	0.78	88.60	89.97
Norouzzadeh	1	0.79	87.76	90.07
Hyperbolic	1	0.81	88.60	91.26
Hyper	2	0.85	90.87	91.44
Mrazek	2	0.75	87.23	89.71
Yang	3	**0.89**	**94.45**	**94.64**
Atto	3	**0.88**	**94.38**	**94.67**

Table 4: Experimental results using Helliger distance

CA shrinkage function	No. of Parameters	F–measure	AUC	Avg. Prec.
Not applied	-	80.36	90.06	91.45
Hard	1	76.20	84.87	85.68
Soft	1	65.46	72.59	73.87
Norouzzadeh	1	52.29	56.98	56.63
Hyperbolic	1	76.50	83.69	85.37
Hyper	2	77.25	85.19	85.76
Mrazek	2	75.90	86.60	88.33
Yang	3	76.51	84.14	83.68
Atto	3	75.03	83.35	82.92

4.3. Fractional delay time embedding of EEG signals into high dimensional phase space

As mentioned in Chapter 2, EEG can be considered to be chaotic, and is usually analyzed using nonlinear methods.

Nonlinear analysis usually depends on the reconstruction of the phase-space geometry of the signal from a small number of observables. The first step is to embed data in a higher dimensional phase space. The embedding method is based on two theorems. First, Whitney Embedding Theorem [Whitney, 1937] implies that $2M + 1$ independent signals measured from a system can be considered as a map from the set of states to the $2M + 1$ dimensional space, thereby reconstructing the phase space. Next, Takens Embedding Theorem [Takens, 1981] proved that instead of $2M + 1$ generic signals, the time-delayed versions of a generic signal would suffice to embed the M-dimensional manifold.

The idea of embedding has been used previously for analysis of EEG data. Athitsos et al. [Athitsos et al., 2008] proposed a method for approximate subsequence matching by using embedding. Embedding maps each database time series into a sequence of vectors, so that every step of every time series is mapped to a vector. Embedding is computed by applying full dynamic time warping between reference objects and each time series. Klonowski et al.[Klonowski et al., 1999] have estimated the embedding dimension of the EEG signals of the same subjects to be between 7 and 11.

Yuan et al. [Yuan et al., 2008] analyzed the embedding dimension of normal and epileptic EEG time series. The study has found that the embedding dimension reflects variation of freedom degree of human brain nonlinear dynamic system during seizure. The embedding dimension changes during seizure and becomes different from that of normal EEG signals: the average value of the embedding dimension of normal EEG signals is 8, while the embedding dimension of the epileptic EEG signals is 17.

We have proposed a method for reconstruction of the high-dimensional phase space of the EEG signal. The novelty of the method, first, is the concept of fractional embedding, which extends linear embedding for fractional time delays. The second novelty is the reconstruction of the fuzzy state vector of the signal which is useful for in-depth signal analysis and identification of signal landmarks [Cortes and Vapnik, 1995].

The aim of binary classification is classification of the members of a given set of objects into two groups on the basis of whether they have some property or not. Let us have two datasets of signal values: a positive dataset $\{X_1^P, X_2^P, \cdots, X_p^P\}$ and a negative dataset $\{X_1^N, X_2^N, \cdots, X_n^N\}$, where $X_i = \{x(t), t \in T\}$ is a time series (signal), representing some real physical process sampled at time t, T is the index space, and p and n are the size of the positive and negative dataset, respectively.

Let us have a label function $\lambda : X \to [0, 1]$, which indicates to which

dataset a time series belongs to, i.e., $\forall X_i^P, \lambda(X_i^P) = 1$ and $\forall X_i^N, \lambda(X_i^N) = 0$

The aim is to perform classification of unknown label time series X_u sampled from X , to find $\lambda(X_u)$.

The proposed method consists of the following stages:

1. **Reconstruction of the high-dimensional phase space of the signal.** The M-dimensional phase space $\mathbb{R}^M$ is reconstructed from a single observable $x(t)$, using the M-dimensional vectors (trajectories)

$$Y = [x(t), x(t+\tau), \cdots x(t+(M-1)\tau)], \tag{45}$$

where τ a free delay (lag) parameter, and M is the dimensionality of the trajectory. We define the reconstruction of the phase space as a function

$$\chi_\tau : X \to Y, \tag{46}$$

where τ is the delay parameter.

The Taken's theorem suggests that theoretically, the choice of this parameter is of no importance. However, in practice it is crucial to choose a good τ value due to noise [Ulbikas et al., 1998]. The classical method of delays allows only for $\tau = \mathbb{N}$ Our novelty is the modification of the classical delay-based method to allow for fractional delay, where $\tau = \mathbb{R}$. The value of $x(t+\tau)$ for a fractional delay τ is calculated using the linear interpolation as in (47).

$$x(t+\tau) = x(\lfloor t+\tau \rfloor) + \frac{x(\lceil t+\tau \rceil) - x(\lfloor t+\tau \rfloor)}{\lceil \tau \rceil - \lfloor \tau \rfloor}(\tau - \lfloor \tau \rfloor). \tag{47}$$

This step allows for dimensionality reduction of original datasets, because usually $M \ll T$. The result of this step is two datasets of signal trajectories: a positive dataset $\{Y_1^P, Y_2^P, \cdots, Y_{p\times M}^P\}$ and a negative dataset $\{Y_1^N, Y_2^N, \cdots, Y_{n\times M}^N\}$.

2. **Selection of positive trajectories from the phase space.** We assume that a positive dataset Y^P is a mixed set of trajectories, which consists of a subset of trajectories that are similar to Y^N as well as of a subset of trajectories that are different from Y^N. The aim is to find a subset of positive trajectories $\hat{Y}^P$ such that the distance between $\hat{Y}_i^P$ and $\hat{Y}_j^P$ is the largest.

We also assume that for any dataset X_i there exists at least one such positive trajectory such as described in (48).

$$\forall \hat{Y}_i^P, \max_j D(\hat{Y}_i^P, \hat{Y}_j^N), \tag{48}$$

where $\hat{Y}_i^P \in \chi(X_k^P)$, $Y_j^N in \chi(X_k^N)$ and $D : Y \times Y \to \mathbb{R}$ is a dissimilarity (distance) function between two trajectories.

Here we use the Euclidean distance, described in (49), as the dissimilarity function. We specify the selection step as a selection operator $\rho : Y \to \hat{Y}$.

$$d_{i,j} = \sqrt{\sum_{k=1}^{M} (y_{ik} - y_{jk})^2} \tag{49}$$

3. **Distance-adaptive sampling of the negative trajectories from the phase space**. The result of the selection step is a positive trajectory dataset $\{\hat{Y}_1^P, \hat{Y}_2^P, \cdots \hat{Y}_M^P\}$, which is significantly smaller than a negative trajectory dataset $\{Y_1^N, Y_2^N, \cdots, Y_n^N \times M\}$. Using such imbalanced datasets for classification is impractical, because the classification results become biased [12]. To balance the dataset, we perform the distance-adaptive sampling of a negative dataset Y^N as described in (50).

$$h_\epsilon(Y_i^N) = |Y_j^N|, D(Y_i^N, Y_j^N) < \epsilon, \tag{50}$$

where h_ϵ is the ϵ-neighborhood count function of trajectories, $i \neq j$ and $Y_i^N, Y_j^N \in \chi_\tau(X_k^N)$.

The sampling probability of Y_i^N is described in (51)

$$p_s(Y_i^N) = 1 - \frac{1}{1 + h_\epsilon(Y_i^N)} \tag{51}$$

The algorithm of the distance-adaptive sampling is as follows:

   (a) Calculate a dissimilarity matrix D, where $D(i, j)$ is a distance (dissimilarity) between Y_i and Y_j.

   (b) Calculate a neighborhood array E, where E_i is the number of neighbors of Y_i which are at most at ϵ distance from Y_j.

   (c) Calculate the probability array P, where P_i is the sampling probability of Y_i.

   (d) Construct a sampled dataset $\hat{Y}^N$, where $Y_i^N \in \hat{Y}^N$ if $p_s(Y_i^N) > p$, and p is a Gaussian distributed random number from the range $\{0, 1\}$.

4. **Classification.** Finally, the trajectory vectors are used as features for classification. For training datasets, we define the label function $\hat{\lambda} : Y \to \{0, 1\{$ and assign $\hat{\lambda}(\hat{Y}_i^P) = 1$, and $\hat{\lambda}(\hat{Y}_j^N) = 0$. The result of training is a support vector model which is used for classification of an unknown label time series X_u as follows:

- Reconstruct the set of trajectories in the phase state $Y^u = \chi_\tau(X_u)$.
- Perform classification of a set of trajectories Y_u and obtain the set of trajectory labels $L^u = \hat{\lambda}(Y_i^u)$, where $Y_i^u \in Y^u$.

5. **Reconstruction of a fuzzy signal state.** The classification result of each trajectory is used to reconstruct the state vector $S(t)$ of the EEG signal as described in (52).

$$S(t) = \frac{1}{M} \sum_{k=t}^{t+M} L_k^u. \tag{52}$$

The results are interpreted in terms of fuzzy logic as a fuzzy probability of the state (positive or negative) of the signal at time t. Finally, the classification label is assigned to X_u as described in (53).

$$\lambda(X_u) = \begin{cases} 1, & \text{if } \max_t S(t) \geq \theta, \\ 0, & \text{if } \max_t S(t) < \theta. \end{cases} \tag{53}$$

where θ is the threshold probability of the state.

For this experiment, datasets Ia, Ib and dataset IV were used. For the dataset IV, down-sampling to 256 Hz was performed before the experiments. All datasets were randomly divided into 5 parts, and 5-fold cross-validation was used to evaluate the classification results.

The results of the fractional time delay embedding of the EEG signal into the high dimensional phase space are shown in Figure 4.1 (dataset Ib, channel 0). The plot shows that positive dataset has a wider distribution of values than the negative dataset (compare Figure 4.1 red with black). Such extreme values (called landmarks [Cortes and Vapnik, 1995]) can be exploited for signal classification.

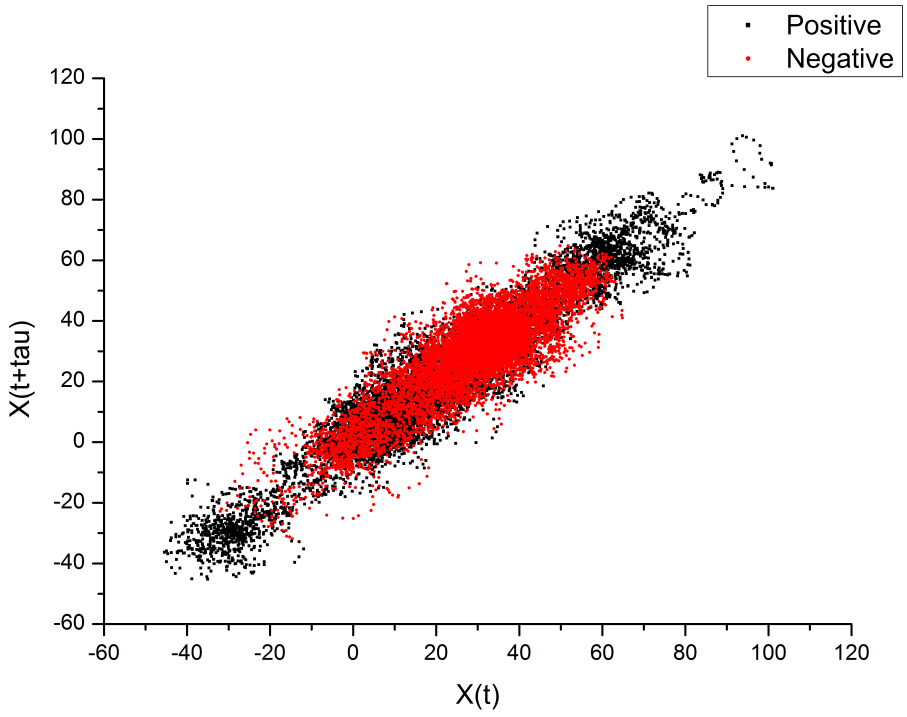


Figure 4.1: Phase space reconstruction of EEG signals (dimension $M = 2$, time delay $\tau = 9.5$): 20 positive signals (black) and 20 negative signals (red)

4.4. The Wave Atom transform

Classifying EEG data requires, firstly, the reduction of its high-dimensional feature space to identify fewer intrinsic feature dimensions relevant to specific mental states of a subject. To perform classification, Artificial Neural Networks (ANN) are a good choice due to their ability to generalize and work well with noisy data [Subasi and Erçelebi, 2005]. However, long training time is a well-known disadvantage of ANN, especially when a BCI system must deal with real-time constraints [Martišius et al., 2013b].

The system classifier needs to be retrained constantly in a real time BCI application. The subject's mental state changes in time, and the classifier must adapt to these changes. Also training must be performed according to the control task at hand, i.e. different video game controls require differently trained classifiers. Training time is particularly important because EEG training datasets are usually large and computationally demanding in order to correctly train the classifier. Data from previous user trials or even other users can be reused.

Presenting a network with a subset of a spectrogram as an input vector is the obvious solution. Such a system has been used effectively to classify EEG data [Addison, 2005]. This technique, however, still has a high computational cost since large numbers of weights must be updated in training. Classical signal transforms (such as FFT or DCT) decompose time series into global sinusoidal components of fixed amplitude. However, the Fourier methods may not give an efficient representation of a signal: as a result of the infinite extent of the Fourier integral, analysis is time averaged. Thus it contains only globally averaged information and so has the potential to obscure transient or location specific features within the signal. This limitation can be partly overcome by introducing a sliding time window of fixed length to localize the analysis in time. This local or short time Fourier transform (STFT) provides a degree of temporal resolution by highlighting changes in spectral response with respect to time. A number of alternative time–frequency methods are now available for signal analysis, yet, we previously wrote in Chapter 4, one of the most popular methods are various wavelet transforms.

Wavelet transforms have uniform temporal resolution for all frequency scales, but resolution is limited by the basic wavelet [Demanet and Ying, 2007]. These transforms are useful for characterizing gradual frequency changes.

Wave Atom Transform (WAT) is a relatively new transform, proposed by Demanet and Ying [Demanet and Ying, 2007]. WAT performs a multi-resolutional analysis of a signal, i.e., decomposes a signal into different frequency sub-bands. Wave atoms are a variant of wavelets that have a sharp frequency localization and offer a sparser expansion for oscillatory functions than wavelets. Wave atoms compose wave-fields as a superposition of highly anisotropic, localized, and multi-scale waveforms and capture coherence of pattern across and

along oscillations.

WAT has been previously used mainly in image processing for image denoising , image watermarking, image hashing, but also for fingerprint recognition [Mohammed et al., 2010], as well as feature extraction [Patwari et al., 2014], dimensionality reduction and numerical analysis [Mohammed et al., 2010]. To our knowledge, wave atoms have not been applied to EEG signal processing. Similar research included ECG (electrocardiogram) [Xu and Zhai, 2011, Dua and Raj, 2012] and MRI (Magnetic Resonance Image) [Rajeesh et al., 2010] data. WAT is a promising approach for EEG processing because of its denoising and feature extraction capabilities, and is particularly useful when the signal has discontinuities and sharp spikes as is in case of EEG.

We expect that WAT coefficients extracted from EEG data samples can retain enough information to permit correct classification, while feature reduction should reduce network training and classification time.

We analyze and perform experimental research on the ANN training functions to find the best combination of training speed and classification accuracy as applied to the WAT coefficients of the EEG data.

Wave atoms are a variant of 2D wavelet packets that retain an isotropic aspect ratio. They are well suited for representing oscillatory patterns in a signal [Mohammed et al., 2010]. A signal is regarded as an oscillatory pattern if it can be described as (54)

$$f(x) = sin(Ng(x))h(x), \tag{54}$$

where x is a coordinate, g and h are the C^∞ scale function, N is a large constant. WAT allows to obtain a sparser solution of the oscillatory pattern signal f than that of other known transformations, as only $O(N)$ wave atom coefficients are sufficient to represent f to some given accuracy [Mohammed et al., 2010], while $O(N^2)$ wavelet coefficients are needed in the same case.

The definition of wave atoms for 1D signal is as follows. Consider $f(x)$ and $\hat{f}(w)$ is a 1D Fourier transform pair, x, w corresponds to the coordinates in time domain and frequency domain. Define wave atoms as $\phi_\mu(x)$, in which $\mu = (j; m; n)$. Then the indexed point $(x_\mu; w_\mu)$ in phase-space is defined in (55).

$$x_\mu = 2^{-j}m \; ; \; \omega_\mu = \pi 2^j n \; . \tag{55}$$

The definition is similar to that of wavelet packets, where j controls the resolution scale, m and n control the location in time and frequency domain.

The elements of frame ϕ_μ are called wave atoms when it satisfies (56).

$$\begin{gathered} |\hat{\phi}_\mu(\omega)| \leq C_M 2^{-j}(1 + 2^{-j}|\omega - \omega_\mu|)^{-M} + C_M 2^{-j}(1 + 2^{-j}|\omega + \omega_\mu|)^{-M}; \\ |\phi_\mu(x)| \leq C_M 2^j (1 + 2^j |x - x_\mu|)^{-M} \; . \end{gathered} \tag{56}$$

for all $M > 0$. The difference between wavelet packets and wave atoms are that wave atoms obey the parabolic scaling wavelength: at scale 2^{-2j} the essential frequency support is of size 2^j while at frequency 2^{2j}, the essential time support is of size 2^j [Demanet and Ying, 2007]. Also, the basis function is different from a conventional wavelet basis function. The frequency domain formula of a basis function is defined as (57).

$$\hat{\Psi}_m^0(\omega) = e^{-i\omega/2}[e^{i\alpha_m} g(\epsilon_m(\omega - \pi\kappa)) + e^{-i\alpha_m} g(\epsilon_{m+1}(\omega + \pi\kappa))] \ . \tag{57}$$

where $\kappa = \left(m + \frac{1}{2}\right)$, $\epsilon_m = (-1)m$ and $\alpha_m = \left(\frac{\pi}{2} + \frac{1}{2}\right)$. The function g is an appropriate real-valued function, compactly supported on an interval of length 2π, and chosen according to (58).

$$\sum_m |\hat{\Psi}_m^0(\omega)|^2 = 1 \ . \tag{58}$$

Data form the BBCI dataset Ia was used for this experiment. To solve the two - class classification problem, we used a strictly feed-forward artificial neural network structure with one input layer, one output layer and one hidden neuron layer, initialized with random values. A tangent sigmoid threshold function was used both in hidden and output layers, since this proved to produce better results. A 10-fold cross validation was performed for every neural network hidden layer size and average accuracy, F-measure and training time was calculated. For comparison, raw EEG data, 50 DCT coefficients and WAT 1st scale coefficients were used as inputs. Training speed was evaluated by an average time needed for one ANN training. A PC running Debian Linux with an AMD Phenom II X4 (3.6 GHz) processor and 8GB RAM was used.

Classification accuracy is presented in Figure 4.2. The picture shows that WAT transform produced better accuracy results than DCT or RAW with all network configurations. Best accuracy of of 90% was obtained using a Bayesian Regularization network training algorithm. This was achieved with just 2 neurons in the hidden network layer. No results for hidden layer sizes of 10 and 15 neurons for RAW EEG data are presented because of high network training time.

Experimental results are summarized in Table 5. Average training time and classification accuracy for a neural network achieving the best results are presented. Training time is dramatically decreased with the use of a reduced training dataset. WAT shows better training time and classification results using Levenberg-Marquardt and Bayesian Regularization functions. The box chart for F-measure results is presented in Figure 4.3. Therefore, WAT has a great potential for application in real-time BCI tasks.

All classification quality results are in line with the best results obtained in the BBCI competition II and other authors, using the same dataset in

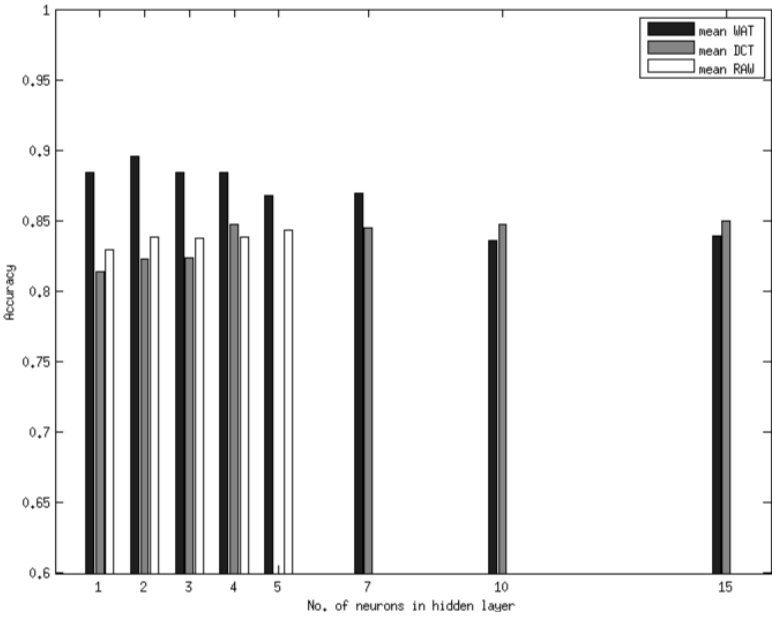


Figure 4.2: Accuracy using Bayesian regularization training of ANN

years, following the competition [Nguyen et al., 2015a, Nguyen et al., 2015b, Xu et al., 2014]. The best accuracy of 90% is achieved using Bayesian Regularization training.

Table 5: Comparison of classification accuracy and network training time

Features	Neurons	Accuracy, %	F-measure	Time, s
RAW	5	84	0.84	11906
DCT	15	85	0.85	40
WAT	2	**90**	**0.90**	**1.1**

Our results show this method to be effective for reducing data size without the loss of important signal information. Data classification was performed using artificial neural networks with various hidden layer sizes and training functions. Results show a dramatic improvement of training speed with a transformed dataset. The best improvement was achieved using Bayesian regularization training. This improvement would mostly benefit real-time BCI applications.

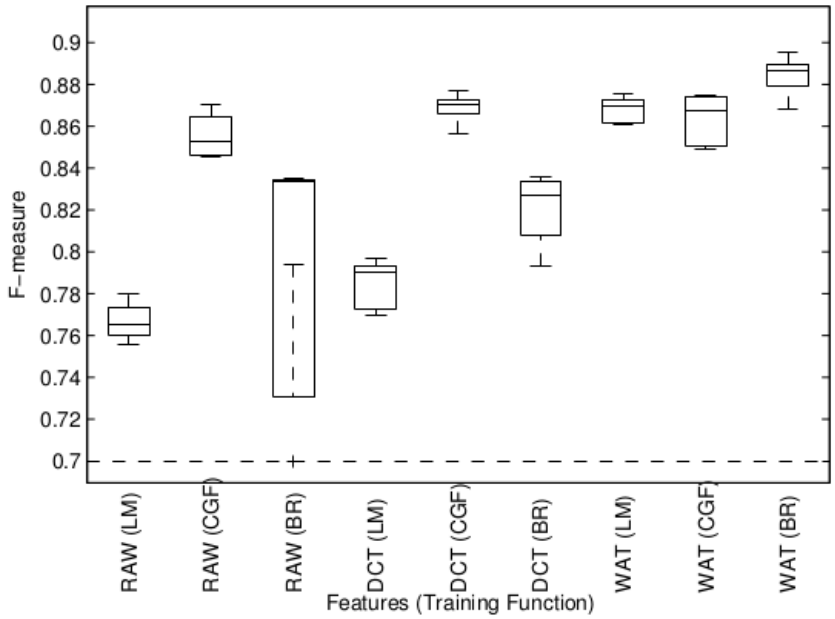


Figure 4.3: F-measure box diagram

4.5. Use of non-linear operators

Recently, the non-linear operators such as the Teager-Kaiser energy operator (TKEO) [Teager, 1980, Kaiser, 1990] have attracted the attention of researchers in the BCI domain. The TKEO so far has been applied in speech recognition, for image enhancement and in EEG data analysis for detecting high frequency oscillations, mental task classification, and sleep spindle detection.

The advantage of the TKEO family of operators over the traditional DSP analysis methods such as Fourier Transform or wavelet analysis is the ability of the TKEO to discover high-frequency low-amplitude components in analyzed data. The TKEO unlike conventional energy takes into account the frequency component of the signal as well as the signal amplitude.

Nonlinear models are systems where either the additivity or the scalability properties do not hold in general, as in (59).

$$H\left[f(t)+g(t)\right] \neq H\left[f(t)\right]+\left[g(t)\right], H\left[\alpha \times f(t)\right] \neq \alpha \times H\left[f(t)\right], \tag{59}$$

where $f(t)$, $g(t)$ are signals, H is a nonlinear model (operator), and α is a constant.

The Teager-Kaiser Energy Operator (TKEO), proposed by Teager [Teager, 1980] and further investigated by Kaiser [Kaiser, 1990], is a special case of nonlinear models. For a continuous real-valued signal $x(t)$, the TKEO $\Psi[x(t)]$ is defined as (60):

$$\Psi[x(t)] = \left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 - x(t)\frac{\mathrm{d}^2x}{\mathrm{d}t^2}. \tag{60}$$

An approximation of the derivatives by one-sample differences provides the definition of the TKEO. For the discrete-time signal, it is defined in (61) [Kaiser, 1990].

$$\Psi[x(n)] = x^2(n) - x(n-1)x(n+1). \tag{61}$$

Moore et al. [Moore et al., 1996] propose a generalization of the Teager operator as 1-D Volterra filter, as shown in (62)

$$\Psi^m[x(n)] = [x(n)]^{2/m} - [x(n-1)x(n+1)]^{1/m}. \tag{62}$$

Tomar et al. [Tomar and Patil, 2008] introduce two generalizations of TKEO. The variable length TKEO (VTEO) is defined as (63)

$$\Psi_i[x(n)] = x^2(n) - x(n-i)x(n+i). \tag{63}$$

The Summed-over Variable length Teager Energy Operator (S-VTEO) is defined as in (64).

$$\zeta_i x[n] = \sum_{k=1}^{i} \Psi_i x[n]. \tag{64}$$

A combination of (62) and (63) operators is proposed in [Kvedalen, 2003] as (65).

$$\Psi_i^m x[n] = x[n]2/m - x[n-i]x[n+i]^{1/m}. \tag{65}$$

The generalization of the continuous TKEO as the higher-order energy operator (HOEO) Ψ_k is proposed in [Maragos and Potamianos, 1995] and defined as (66).

$$\Psi_k[x(t)] = \frac{\mathrm{d}x}{\mathrm{d}t} \times \frac{\mathrm{d}^{k-1}x}{\mathrm{d}t^{k-1}} - x(t)\frac{\mathrm{d}^k x}{\mathrm{d}t^k}. \tag{66}$$

For discrete-time series, the HOEO can be rewritten as the discrete energy operator (DEO), as in (67) [Maragos and Potamianos, 1995]:

$$\Psi_{km}[x(t)] = x[n]x[n+k] - x[n-m]x[n+k+m]. \tag{67}$$

In the general case, the TKEO operator can be generalized to the Homogeneous Multivariate Polynomial Operator (HMPO) $\Psi_m^k[x(t)]$, where the 2nd order HMPO is defined as in (68):

$$\Phi_k^2[x(n)] = \sum_{i=-z}^{z} \sum_{j=-z}^{z} A_{ij} x[n+i] x[n+j], \tag{68}$$

where $z = \lfloor m/2 \rfloor$, and A is the coefficient matrix. The 3rd order HMPO is defined as (69):

$$\Phi_k^3[x(n)] = \sum_{i=-z}^{z} \sum_{j=-z}^{z} \sum_{k=-z}^{z} A_{ijk} x[n+i] x[n+j] x[n+k]. \tag{69}$$

For example, TKEO, defined earlier in (60) can be written as (70).

$$\Psi[x[n]] = \Phi_3^2[x[n]], \tag{70}$$

where:

$$A = \begin{pmatrix} 0 & 0 & -1/2 \\ 0 & 1 & 0 \\ -1/2 & 0 & 0 \end{pmatrix}. \tag{71}$$

Such operator also can be seen as a special case of the 2D Volterra system as noted by Kvedalen [Kvedalen, 2003]. The properties of the operator are as follows:

- Symmetry. Reversing the signal in time does not change the resulting value;
- Robustness. The operator is robust even if the signal passes through zero, i.e. $x(n) = 0$, there is no division operation.
- Complexity. Complexity of the operator is $O(m^k)$.

For experiments, Data set Ia was used. The following nonlinear operators were applied to the raw EEG data: TKEO (61), TKEO-Volterra (62), VTEO (63), VTEO-Volterra (65), DEO (67) and the proposed HMPO (69), where the non-zero elements of the 3D matrix are: $a_{-1,-1,-1} = 1$, $a_{-1,-1,0} = 3$, $a_{-1,-1,1} = -1$, $a_{-1,0,0} = 2$, $a_{-1,0,1} = -2$. The matrix values were set based on the results of grid-based search using all possible combinations of integer numbers from the range $[-3, 3]$.

The classification of the data was performed using the SVMPerf [Joachims, 2005] implementation of Support Vector Machine a with linear kernel. Kernel parameters were optimized using the Nelder-Mead algorithm based

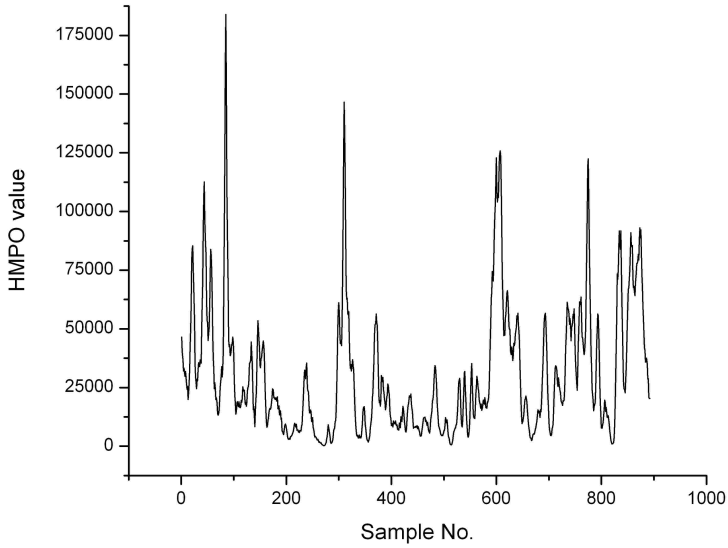
Table 6: Experimental results using different polynomial operators

Operator used	Classification metric				
	Acc., %	Prec.	Rec.	F	AUC
None (original EEG data)	78	0.7462	0.8501	0.7891	0.9018
TKEO (61)	47	0.4805	0.7527	0.5849	0.4670
TKEO-Volterra (62)	59	0.5657	0.7889	0.6583	0.5428
VTEO (63)	56	0.5415	0.8051	0.6466	0.6500
VTEO-Volterra (65)	48	0.4207	0.5330	0.4214	0.1381
DEO(67)	53	0.5160	0.7202	0.6063	0.5851
HMPO (69)	**83**	**0.8042**	**0.8727**	**0.8349**	**0.8450**

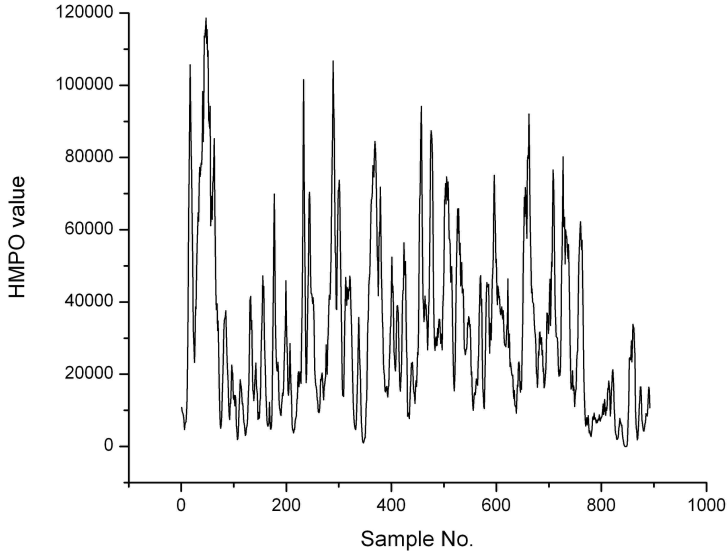
method described in [Nelder and Mead, 1965]. The results of experiments are summarized in Table 6 (the best results in each category are shown in bold).

Evaluating the experimental results, we can claim that the 3^{rd} order nonlinear operators such as the proposed HMPO can demonstrate better results for feature identification of the EEG data than traditional 2^{nd} order operators such as TKEO or its generalizations as demonstrated by Table 6. By visually inspecting the graphs of positive and negative series in Figure 4.4, one can see that the HMPO operator allows for better identification of significant features (slow cortical potential signals of 1 Hz frequency).

The proposed operator can be used for developing new EEG signal processing algorithms, which can be used in Brain-Computer Interface applications, i.e., for robot control in the noisy environment.



(a) positive class instance



(b) negative class instance

Figure 4.4: Samples of EEG data after application of HMPO operator

4.6. Real-Time Training of Voted Perceptron

The main requirement for the BCI systems is to be real-time, i.e. to be able to operate and guarantee response within strict time constraints [Ben-Ari, 1990], which are determined by the operated electronic device. The most time-consuming computations in the BCI systems are usually performed by the classifier, which extracts control classes from the EEG data. However, most of the existing classification algorithms are not suitable for large-scale real-time computation due to their computational complexity.

Current fast BCI classifier training include:

1. Feature dimensionality reduction, e.g., using a portion of the Discrete Cosine Transform (DCT) coefficients of the signal for classification thus reducing the number of features significantly [Birvinskas et al., 2012]. Feature extraction methods such as Principal Component Analysis and Independent Component Analysis can also be used to reduce the dimensionality of the EEG signal [Álvarez Estévez et al., 2011].

2. The non-linear operators, such as in 4.5 or the autocorrelation function [Wright et al., 1990] can be applied to the EEG signal to decrease the computational complexity of the classifier.

3. Classification algorithms specifically developed for data stream mining, e.g., the MOA platform [Bifet et al., 2010] for online learning from data streams, and Very Fast Decision Tree algorithm [Gama and Rodrigues, 2009] designed to meet the constraints of the data stream.

4. Using parallel computations to reduce computation time, e.g., novel ANN architectures adapted for parallel computation such as Parallel Layer Perceptrons [Caminhas et al., 2003] and distributed training strategies for reducing training times when large computing clusters are available [McDonald et al., 2010]. Distributed cluster computation for many batch training algorithms has been examined in [Chu et al., 2006].

The novelty of this research is a modification to the training algorithm of a Voted Perceptron [Freund and Shapire, 1999] to adapt it for real-time tasks in BCI applications.

A Voted Perceptron is an algorithm for linear classification, which combines the Rosenblatt's perceptron algorithm with Helmbold and Warmuth's leave-one-out method [Freund and Shapire, 1999]. All weight vectors encountered during the learning process vote on a prediction. A measure of correctness of a weight vector, based upon the number of successive trials in which it correctly classified instances, can be used as the number of votes given to the

weight vector. The output of the Voted Perceptron can be calculated by (72):

$$y_i = sgn\left\{\sum_{n=0}^{N} c_n sgn(w_n x_{i,n})\right\}, \tag{72}$$

where $x_{i,n}$ are inputs, w_n are weights, v_i is the prediction vector, y_i is the predicted class label, d_i is the desired label and e_i is the error. The result of training is a collection of linear separators $\{w_1, w_2, \cdots, w_n\}$ along with the survival time w_n, which is a measure of the reliability of w_n. The architecture of the Voted Perceptron is presented in Figure 4.5.

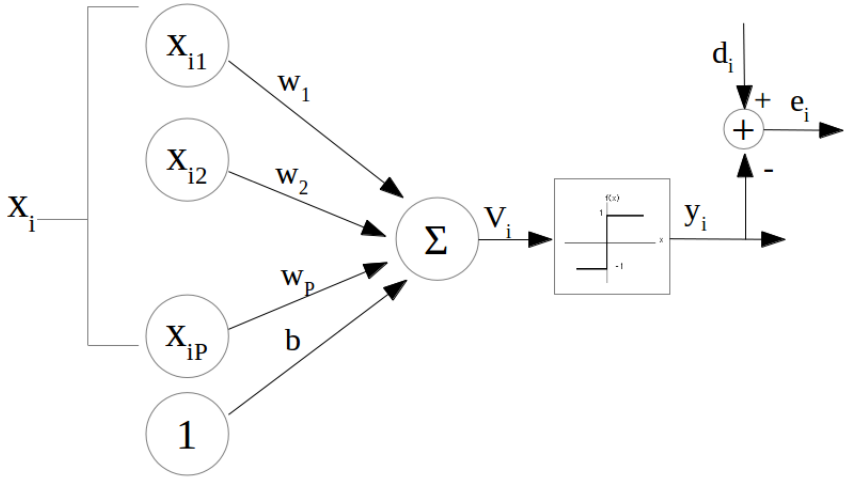


Figure 4.5: Architecture of a Voted Perceptron

The shortcoming of the Voted Perceptron is that its training time is usually unbounded and depends on the size and complexity of training data. According to [Witte, 2005], "the algorithm iterates until a perfect solution has been found, but it will only work properly if the data is linearly separable". Once a misclassified instance is found during training, the hyperplane is modified and the miscategorised instances are reclassified. If the data is linearly separable, the number of iterations is finite. Otherwise, the algorithm will loop infinitely; therefore the maximum number of iterations must be specified. To make Voted Perceptron suitable for real-time BCI applications, we propose the following modification of the training algorithm (see Figure 4.6). The algorithm measures the time elapsed from the beginning of the training and cuts the training procedure as soon as the time bound is reached. For the implementation of the Real-Time Voted Perceptron, we used the open source implementation of the Voted Perceptron algorithm from the Weka package. Weka [Hall et al., 2009] is a suite of machine learning software written in Java.

```
procedure trainClassifier
begin
  let startTime be current time
  Read and filter data
  Randomize training data

  /** Compute perceptrons */
  index := 0;
  i := 0;
  while(true) begin
    instance := 0;
    while(true) begin
      prediction = makePrediction(index, instance);
      if (prediction == classValueOf(instance))
        increaseNeuronWeight(index);
      else begin
        setNeuronWeight(index, i, classValue);
        index := index + 1;
        increaseNeuronWeight(index);
      end;
      let currentTime be be current time
      elapsedTime = currentTime - startTime
      if (elapsedTime >= T)
        finish procedure trainClassifier;
      instance = :instance + 1;
    end;
    i := i + 1;
  end;
end;
```

Figure 4.6: Modified training algorithm of Real-Time Voted Perceptron

To improve the training time, the computational load of the training algorithm must be reduced. To achieve this we propose a dataset transformation to reduce the size of the data. The following transformations can be used to decrease the training time:

1. Down-sampling (or "sub-sampling") reduces the sampling rate of a signal by removing some of the samples of the signal. Down-sampling by n reduces the number of the signal samples by the factor of n.

2. Row sampling is used to obtain a randomly selected subset of an input dataset by selecting only a specified number of the dataset input rows. The rows in the dataset are selected randomly, to make the resultant dataset representative of the original dataset.

For our experiments, Datasets Ia, Ib and dataset IV were used. All datasets were randomly divided into 5 parts, and 5-fold cross-validation was used to evaluate the classification results.

We have performed our experiments on an Intel Core 2 Duo CPU E6550 2.33 GHz PC with 2GB of RAM running Microsoft Windows XP Professional OS.

We performed experiments with the original datasets as well as applied data down-sampling and row sampling techniques. Both original Voted Perceptron and the proposed Real-Time Voted Perceptron training algorithms were used.

To evaluate the quality of classification, we used Cohen's Kappa statistic. The training results are summarized in Figures 4.7-4.16.

Figures 4.7-4.9 for different datasets shows that the number of neurons in a Voted Perceptron depends both on the down-sampling rate as well as on the time bounding value. In the time-unbounded training, the number of neurons fluctuates around 200 for all down-sampling rates. In the time-bounded training, where the time bound was set to 500 ms, the number of neurons in the neural network increases linearly as the down-sampling rate increases until it reaches the same size as in the time-unbounded training. The result can be explained by the fact that down-sampling significantly reduced computational load and, consequently, training time. As the training time in the unbounded training decreases to the bound value, the network sizes also converge.

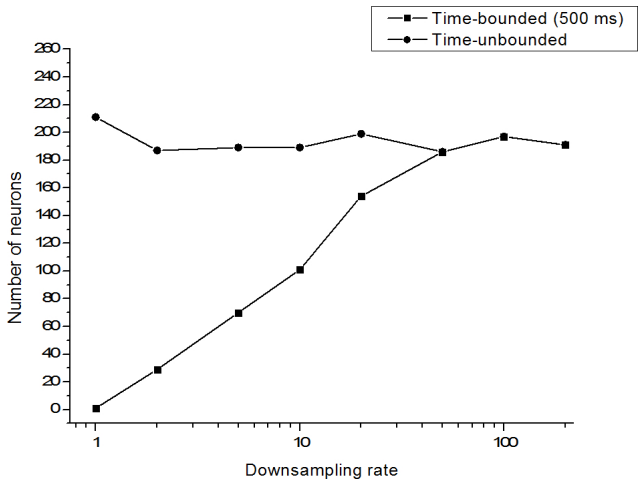


Figure 4.7: Classification quality (κ statistic) after row sampling of Dataset Ia

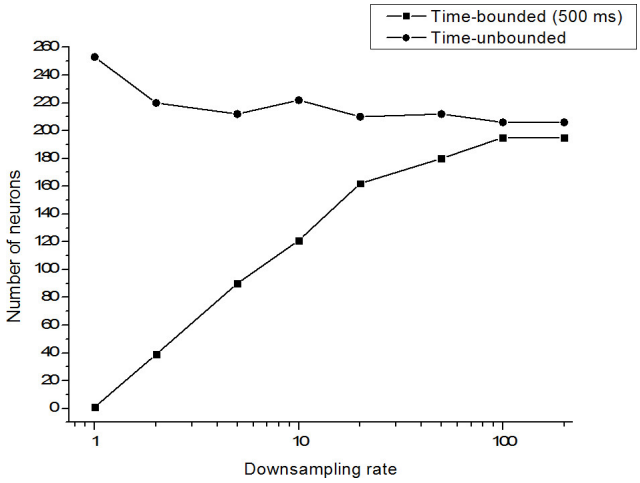


Figure 4.8: Classification quality (κ statistic) after row sampling of Dataset Ib

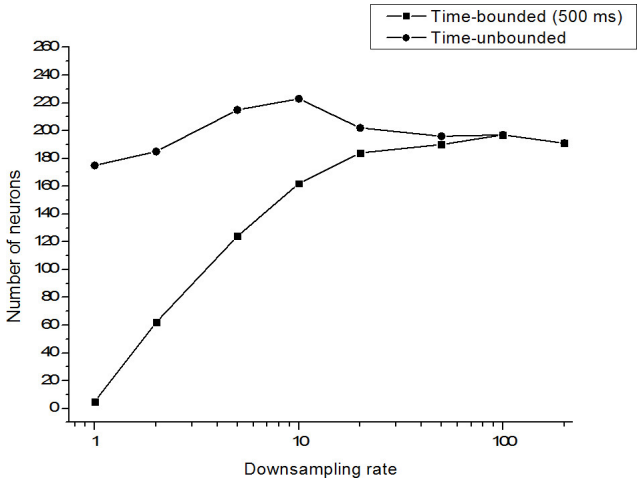


Figure 4.9: Classification quality (κ statistic) after row sampling of Dataset IV

Figure 4.10 shows that while using the original untransformed dataset, network training lasted up to 35 s, yet down-sampling the signal by a factor of 20 or higher can reduce this training time up to the 500 ms or less, which is well suited for real-life BCI use.

Figures 4.11-4.13 show the quality of classification evaluated by using the Kappa statistic (higher values are better) for datasets Ia, Ib and IV accordingly. In the time-unbounded training, classification results depend little on the down-sampling rate (if down-sampling is less than 200). The results show

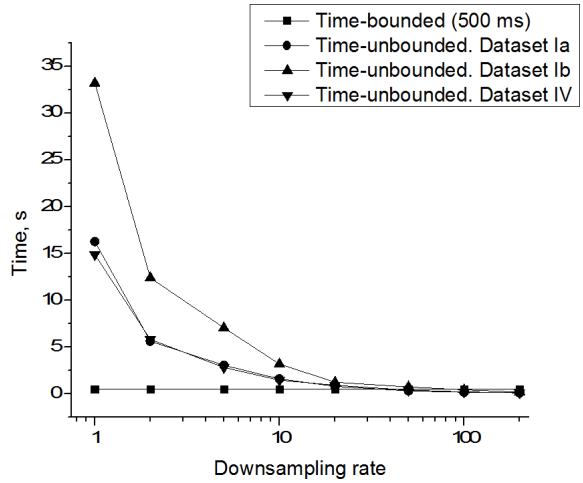


Figure 4.10: Training time of Voted Perceptron

that the EEG signals can be significantly down-sampled without significant loss of classification quality. In the time-bounded training, the training failed on all original datasets within the specified time-bound (500 ms). While with higher down-sampling rates, the time-bounded training produced networks that achieve the same or better classification results than the network produced after time-unbounded training. The explanation is as follows: time-bounded training may prevent the network from over-fitting the data, i.e., the situation when a network demonstrates good results on training data, but shows poor results on test data.

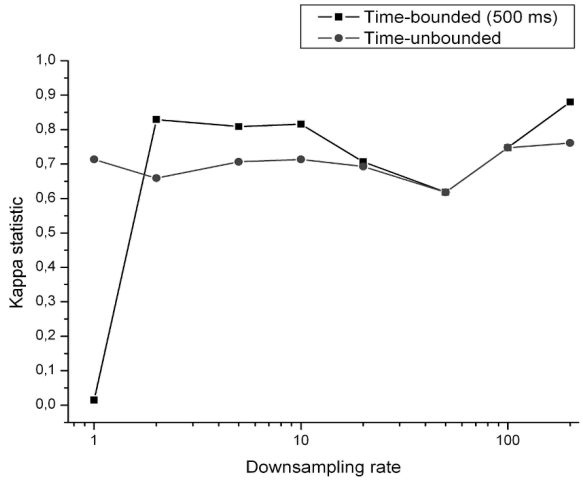


Figure 4.11: Classification quality (κ statistic) after down-sampling - Dataset Ia

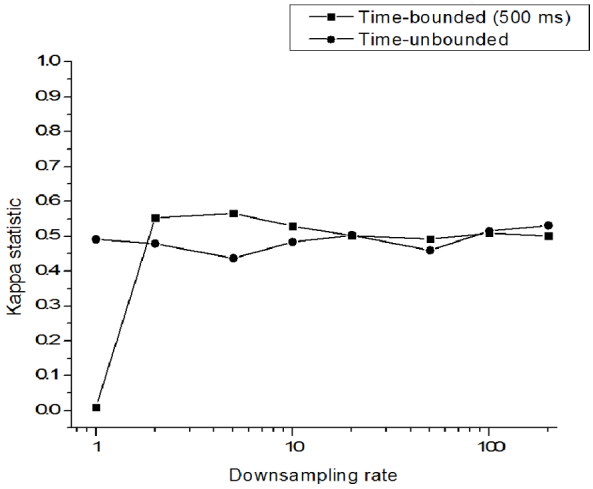


Figure 4.12: Classification quality (κ statistic) after down-sampling - Dataset Ib

Figures 4.14-4.16 show classification quality of the Voted Perceptron after applying the Real-Time Voted Perceptron training algorithm for 200 ms and 500 ms, respectively. For small row sampling values, the real-time training fails. Larger row sampling rate allows for better classification results, which, however, are still worse than after down-sampling. After reaching maximal value at 10, for larger row sampling rates the classification quality diminishes significantly.

4.7. Proposed system design – SSVEP based real-time BCI gaming system

In Chapter 4, we have studied and proposed feature extraction algorithms based on wavelet transforms, thresholding, nonlinear operators and presented an algorithm for time-bound classifier training. We have shown that all these methods could effectively be used in modern BCI systems. In this chapter, we propose a method, based on the Wave Atom transform, to design a BCI control system.

Several other factors have to be taken into account when designing a system prototype. To design an end-user friendly system, which could be used in everyday activities, i.e. wheelchair or mouse cursor control, the system should allow its users to send commands at any time. Such a system must analyze EEG signals continuously, and determine, weather the user is intending to issue control commands to the system i.e is in the Control State (CS); or if he is in a No Control state (NC), indicating that no control commands are issued. If the system detects the user's CS state, it must then decide, which control command is being issued. We take this into account, when designing our BCI system.

Another important thing to consider is the cost, associated with the sys-

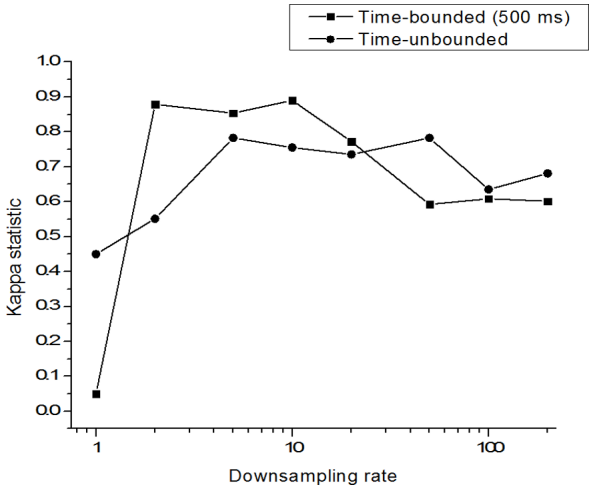


Figure 4.13: Classification quality (κ statistic) after down-sampling - Dataset IV

tem. Although present results in the research field are promising, widespread use of BCI is limited due to the cost and the complexity of equipment. The cost of purchasing professional EEG recording equipment can lie between €20000 and €40000 and has become an obstacle for these systems to become popular commercially. Another obstacle is the requirement of expert knowledge to set up and use these systems.

Recent technological advantages have enabled the development of low cost BCI devices that are aimed at the mass market. The Neurosky Mindwave [NeuroSky, 2014] and the Emotiv EPOC headset [Emotiv, 2014] are examples of such low cost devices.

The goal of our system is to introduce BCI technology as a gaming controller option, which can require less EEG quality and present low risk interactions. By using such low cost devices, aimed at consumers rather than scientists and medics, in the system, we sacrifice performance for price and comfort of the system user.

The EPOC headset, designed by Emotiv inc. [Emotiv, 2014], has been selected as the basis for our system. Basic information about this headset is presented in Section 4.7.1. Later in Section 4.7.2 we present software tools, used for our system design, followed by experimental setup description in Section 4.7.3 and classifier results for system performance in Section 4.7.4. Section 4.7.5 presents the online application of the system.

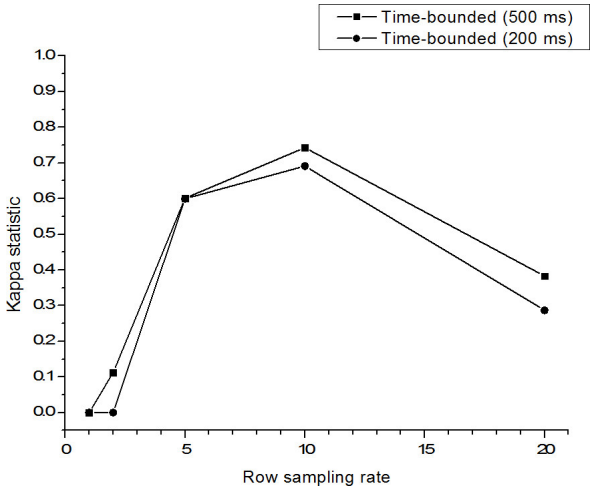


Figure 4.14: Classification quality (κ-statistic) results after row sampling. Dataset Ia

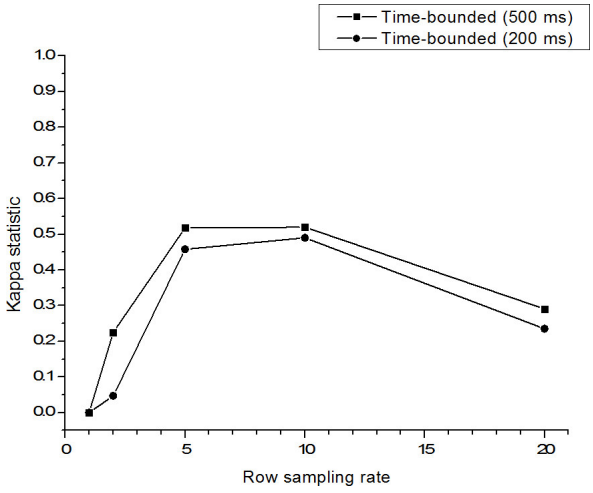


Figure 4.15: Classification quality (κ-statistic) results after row sampling. Dataset Ib

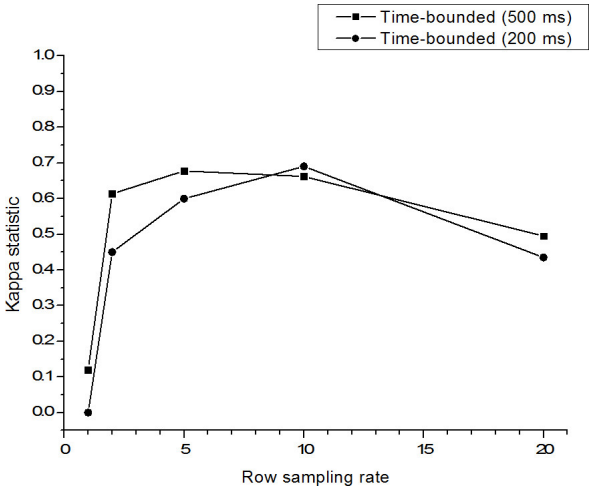


Figure 4.16: Classification quality (κ-statistic) results after row sampling. Dataset IV

4.7.1. The Emotiv EPOC headset

The Emotiv EPOC, shown in Figure 4.17 contains 14 electrodes and 2 reference electrodes, placed in the 10-10 system and labeled as shown in 4.18.

Figure 4.17: Emotive Epoc device [Emotiv, 2014]

The headset is designed as a video game accessory where developers are interested in using the device as a controller. The product chosen for this project was the Research Edition. This provides both the interface for programming with the headset and access to raw EEG data.

The headset transmits encrypted data wirelessly. The wireless chip is proprietary and operates in the same frequency range as 802.11 (2.4Ghz). The internal sampling rate of the device is 2048 Hz. The data is then downsampled to 128 Hz before becoming available to the system for capturing the EEG signals. The captured data contains values for each of the 14 electrodes on the EPOC headset.

There are many advantages to using the Emotiv headset over other BCI and EEG devices. Many BCI devices are restrictive due to wiring. the Emotiv headset, however, is wireless and therefore offers free range of motion allowing for easy transport and setup, which is very important in an everyday use setting. Another advantage is that the EPOC does not require conductive gel for electrodes, making it easier to put on and use. Users do not have to wash their hair after using the headset. The main benefit is that it is relatively inexpensive.

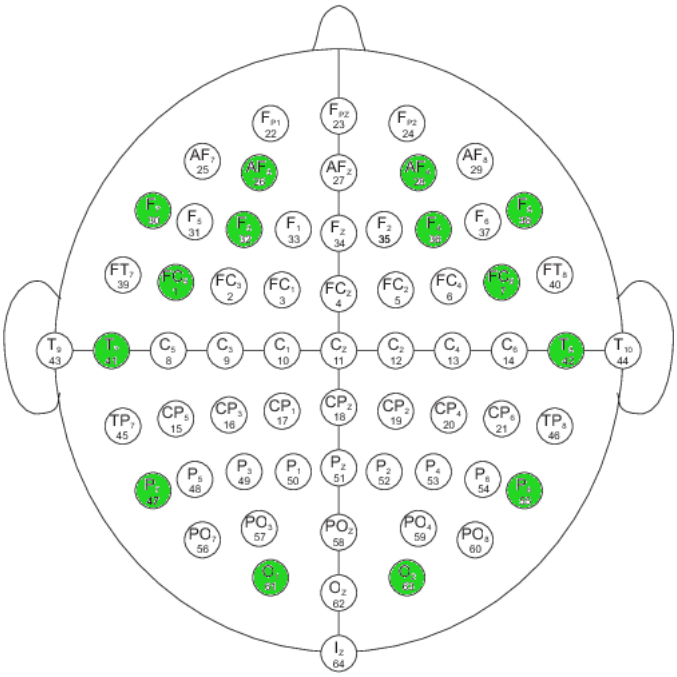


Figure 4.18: Emotive EPOC scalp sensor layout

There are several disadvantages to the EPOC headset as well: it only uses 14 sensors, while many medical grade devices use up to four times that amount. This results in less data coming in from the brain. Additionally, more powerful devices have a sample rate of up to 1000Hz, as opposed to the 128Hz that EPOC runs at. The Emotiv EPOC provides worse signal quality than, for instance, the Biosemi ActiveTwo research system. Since the EPOC headset is not intended for finer signal detection, the electrodes pick up a lot of noise. Several techniques, presented in the previous chapters and tested on offline data, are proposed here to increase the SNR, such as bandpass filtering and averaging.

Researchers have applied the Emotiv neuroheadset's technology in a variety of ways: [Liu et al., 2012] compared the EPOC device to a g.USBamp device in a SSVEP system with good results. It is also used in other paradigms, such as the P300-based system, developed by [Duvinage et al., 2012].

4.7.2. Software tools

This section describes software tools, used for experimental setup and system development.

4.7.2.1 MATLAB

The MATLAB programming language and environment was chosen to implement the system because of its power and ease of use. It is very easy to quickly test various implementation details. MATLAB also has toolboxes which allow additional functionality. Two toolboxes in particular were used for this demonstration. The signal processing toolbox was used extensively for analyzing the EEG data from the headset.

Also, the Emotiv Software Development Kit (EDK) was used for interfacing with the EPOC. It is accessible from a variety of programming languages. It is primarily written in C, but the company also provides wrappers and/or example code for accessing the Application Programming Interface (API) in C++, C#, Java, and MATLAB. MATLAB provides methods for calling functions in C code which allows for straightforward access to the EDK's API.

4.7.2.2 OpenViBE

OpenViBE [Renard et al., 2010] is an open source graphical programming language used to design BCI applications.

The aim of the OpenViBE is to provide open-source software for brain-computer interfaces. Key features of this software are its modularity, high performance, real-time data acquisition and feedback capabilities, compatibility with various hardware devices and multiple scripting language support. It can be used to acquire, filter, process, classify and visualize brain signals in real time. The main OpenViBE application fields are medical (assistance to disabled people, real-time biofeedback, neurofeedback, real-time diagnosis), multimedia (virtual reality, video games), robotics and all other application fields related to brain-computer interfaces and real-time neurosciences.

4.7.2.3 BBCI toolbox

The BBCI Matlab Toolbox [Blankertz et al., 2010] offers a flexible framework for offline and online processing and analysis of neuroimaging (EEG and NIRS) data. The software is available as open source, but is not extensively tested.

It is a collection of MATLAB functions to allow setup for different kinds of BCI experiments, signal processing, data visualization and classification. Some functions of this toolbox have been used for data visualization in our work.

4.7.3. Experimental setup

The objective of this experiment was to develop a system that utilizes brain activity to offer control within a real time environment in order to evaluate signal processing algorithms. A 3 class self paced BCI design, with a NC state

was chosen, as this system setup could easily be adapted for wheelchair or mouse cursor control.

The system is based on the OpenViBE platform and is comprised of 5 individual scenarios, each performed in sequence. The EEG data is recorded using the Emotiv EPOC headset. Since the headset does not have any sensors over the motor cortex, obtaining even moderate results with the Motor Imagery approach is very unlikely. Since the sensors cover the occipital and parietal cortex reasonably, the SSVEP in the multiple visual stimuli selective attention paradigm was chosen for the experiment, due to its well publicized success and limited subject training requirements.

4.7.3.1 Preparation

The preparation for online system usage begins with ensuring that the EEG signal is being recorded properly. In order to obtain a good EEG signal, the sensors have to be moistened with saline solution and have good contact with the skin. The reference electrodes also have to be positioned properly in order to baseline noise-free EEG signals. The first scenario has been created to test, whether the acquisition devices are working properly. The scenario, presented in Figure 4.19 is used to test data acquisition and signal quality.

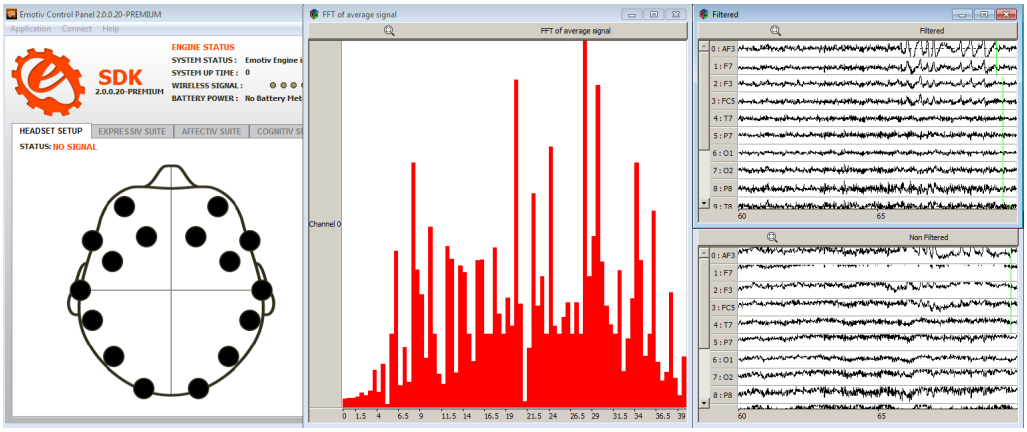


Figure 4.19: Preparation user interface

It begins by initializing the headset driver and displaying the users own EEG signals and sensor contact information on the screen. The sensor contact quality is measured by a proprietary algorithm, provided by Emotiv. This allows for individual sensor contact quality evaluation. The signals of interest, in the case of SSVEP – O1, O2, P7 and P8 can be reviewed individually, as well as the output of the two spatial filters used. Figure 4.20 indicates the electrode montage locations used for this experiment.

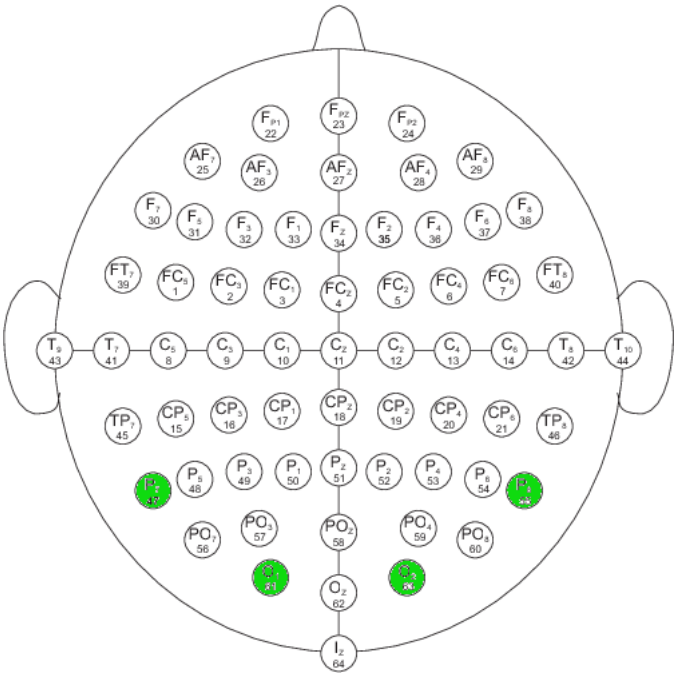


Figure 4.20: Sensor placement in the SSVEP paradigm setup

All channels are averaged and the frequency spectrum is calculated by performing an FFT on the averaged signal. This is also displayed to the user, to check for signs of 50 Hz mains noise or other artifacts in the frequency domain. The users can observe how different facial expressions, clenching their teeth or blinking influences the signal.

This is aimed at showing them the need to be relaxed and avoid face muscle movement or eye blinking. In order to obtain an FFT, the scenario performs the preprocessing of the signal. This is performed by selecting the channels which will be used for data recording. The signal from these sensors are then averaged and band pass filtering of the 6-40 Hz band is performed. The signal is then split into epocs of 2 s, with a 0.5 s interval. An average signal value is obtained by averaging 4 epocs, and an FFT is then performed to visualize the different frequency bands. The user interface window for data aquisition testing is shown in Figure 4.21

Next, some constant settings have to be presented to the system, such as the screen refresh rate of the display, target stimulus color, stimulation frequencies. This only needs to be done once for a particular experimental setup.

Visual user stimulation can be performed by using light emitting diode (LED) or an LCD computer. The LEDs need extra hardware to generate a constant frequency. For the purposes of this experiment, we prefer to use LCD monitors. The drawback of using a monitor is that a stimulus frequency is

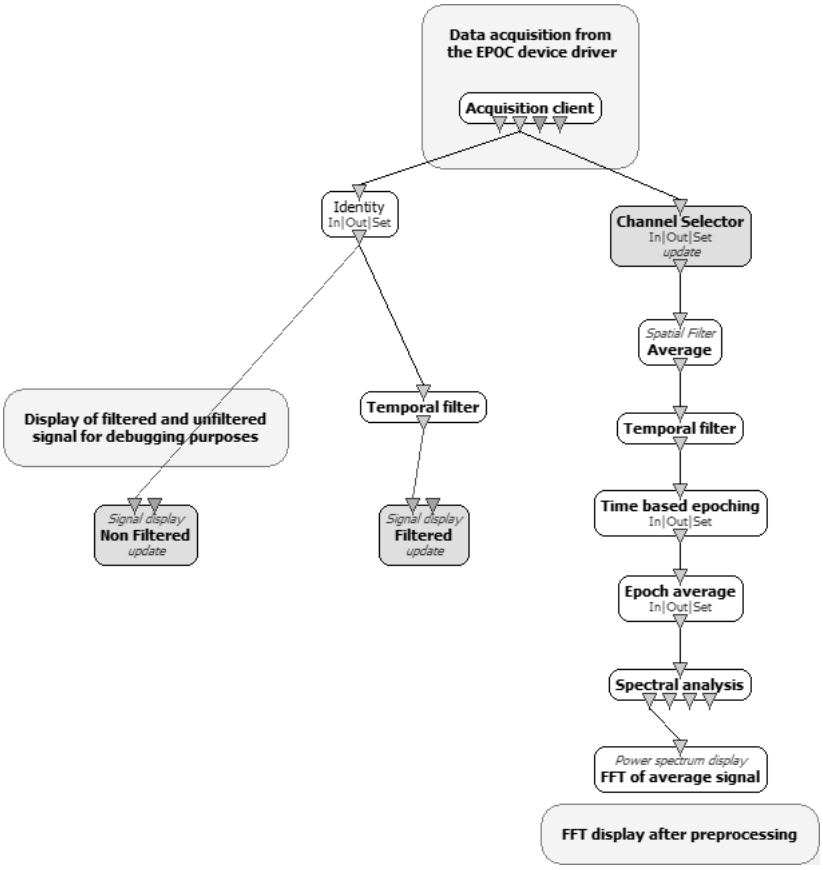


Figure 4.21: Scenario for testing the acquisition and sensor placement

limited by the refresh rate. The refresh rate should be multiple times of the stimulus frequencies i.e., for a monitor with 60 Hz refresh rate, 6.67 Hz, 7.5 Hz, 8.57 Hz, 10 Hz, 12 Hz, 15 Hz and 20 Hz are usually used. When choosing stimulus frequencies it is also important, that a frequency is not a harmonic of another chosen frequency (e.g. 7.5 Hz and 15 Hz). An SSVEP response can trigger a large amplitude response not only in the main frequency, but also in the harmonic frequency, leading to miss-classification. In a 60 Hz refresh rate monitor, for a 10 Hz flicker, it reverses the target color, usually between some light and dark combination, to produce a flicker, every three frames; for 12 Hz flicker, three frames of dark followed by two frames of light color are displayed. Therefore, the sequences of certain two frequencies could be combined to get three frequencies with a varying number of frames in each cycle (e.g. 10 Hz and 12 Hz produce 10.5 Hz, 11 Hz and 11.5 Hz).

A subject, conducting the experiment while wearing the Emotiv EPOC headset can be seen in Figure 4.22.

Figure 4.22: A test subject, wearing the Emotiv EPOC device during an experiment

The EPOC headset has a sampling rate of 128 Hz, and therefore has low resolution at higher frequencies. SSVEP experiments often include stimulation frequencies of up to 60Hz, but these should be avoided while using the EPOC.

Experiments with different frequencies showed that best results for a three class BCI were obtained by using 30 Hz, 20 Hz and 12 Hz. Therefore, these were chosen for the final experimental setup. The color for the flickering targets was chosen as a combination of white and black. The study [Cao et al., 2012] analyzed how different colors of the targets influence classification quality. For our experiments, the white – black color combination was chosen, since it gives the highest contrast.

4.7.3.2 Training data acquisition

In order to perform classifier training, some data needs to be gathered and analyzed. This is done in a supervised learning approach, where the user is presented with an LCD display, containing 3 blinking targets on a black background and a yellow arrow. On cue, the targets start blinking at different

frequencies. This is presented in Figure 4.23.

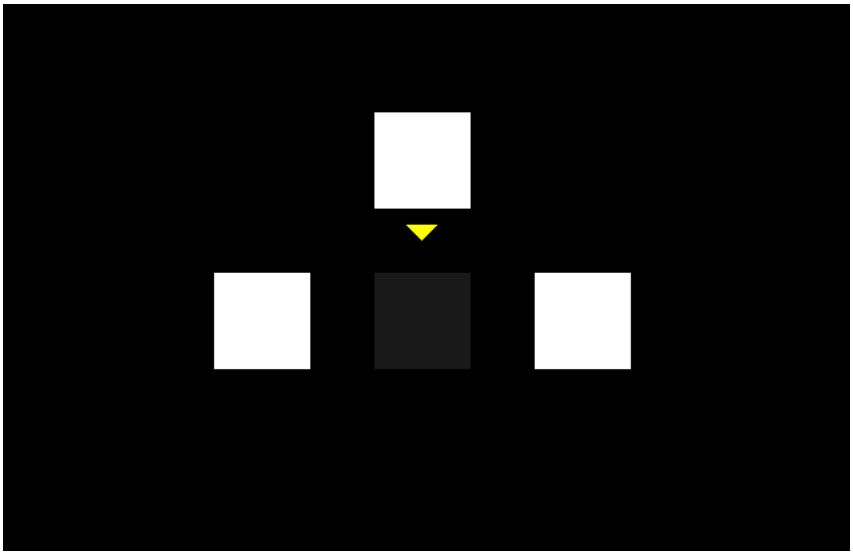


Figure 4.23: Traning data collection interface

The training data acquisition procedure is performed a number of times. Desirably – as many as possible, to obtain the training data for each of the three classes. Since the training session is time-bound by the system, it requires a lot of attention and concentration from the user, and due to fatigue, has to be limited to a number of trials. In this experimental setup, a number of 20 trials per class was chosen, totaling in 60 trials for a single dataset. Classifier training was achieved by gathering 4 EEG data sets of SSVEP data, acquired from 2 healthy subjects (28 year old). Subjects had very few or no previous experience in BCI. During the experiments, they were asked to focus attention on targets, blinking in a defined frequency. A session was composed of 20 trials of each of the three classes (LEFT, RIGHT and CENTER), arranged in a random order.

The timing of the sessions was organized accordingly: in our protocol, the trial lasted 10 s. First, a yellow arrow is displayed for 1 s, indicating the target, on which the user must focus his attention. From second 1 to second 6, the trial entered stimulation phase. In this phase all thee targets start blinking in their corresponding frequencies. The users are specifically ordered not to move the head, relax face muscles and not to blink during this phase. Stimulation is then followed by a 4 s resting period, at which the user is allowed to rest his gaze, blink or move the head. The EEG data from this period is not used for classification. This is illustrated in Figure 4.24.

Training data acquisition is performed in the scenario, presented in Figure 4.25.

The recorded EEG data, together with marked events, such as class labels

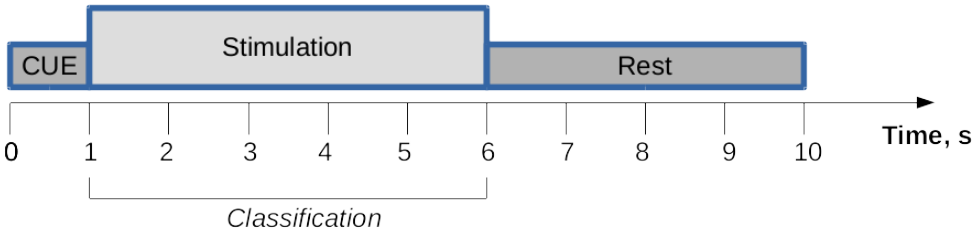


Figure 4.24: Timing of a single SSVEP trial

for each trial are saved on the computer. Next, only relevant channels, in this case O1, O2, P7 and P8 are selected for analysis. Preprocessing steps are then performed in order to denoise the signal and extract relevant information features for the classifier. First, data is split into thee groups, according to their corresponding class label - LEFT, RIGHT and CENTER accordingly. This is done so that a binary classifier could then distinguish, weather a trial belongs to a certain class or not, by using the "one versus all" criteria. This allows for the NC class, where output is false for all three classifiers. Next, temporal and spatial filtering is applied to each of the three groups. Specifically, each group of signals is band-pass filtered around the target frequency of interest: for the LEFT class – 29,5 – 30.5 Hz; CENTER – 19.5–20.5 Hz; RIGHT –11.5–12.5 Hz. This is done, using a 4-th order butterworth filter.

To extract features, we first segment the signal, extracting the 5s long stimulation period from the trial, since only this portion of the signal carries relevant information. Next, each segment is further divided into epochs of 1s every 0.2s, which provides 80% overlap between neighboring epochs. Then, the Wave Atom transform coefficients are obtained for every epoch, and a feature vector is aggregated. As such, 25 feature vectors are extracted for every trial. They are then used for classifier training.

In order to compare this approach to other successful BCI feature extraction techniques, another feature extraction algorithm is also performed on the same data. Features are extracted by training an adaptive CSP filter, then band-pass filtering the signal around the target frequency, as described above, and then performing band power calculation. CSP filters are described in detail in Section 3.2.1.3. Band power values are then used as features to train a classifier. This scenario is presented in Figure 4.26.

The classifier training procedure is done in off-line mode, and can therefore be repeated as many times as needed. Data can also be used in further training sessions or on other users, to reduce the number of user training trials.

Classification results have been evaluated with different classifiers - an SVM with a linear kernel, an SVM with a radial basis kernel and a sLDA classifier. These are covered in detail in Section 3.6.1.

By learning from the feature vectors, the classifier can learn the differences

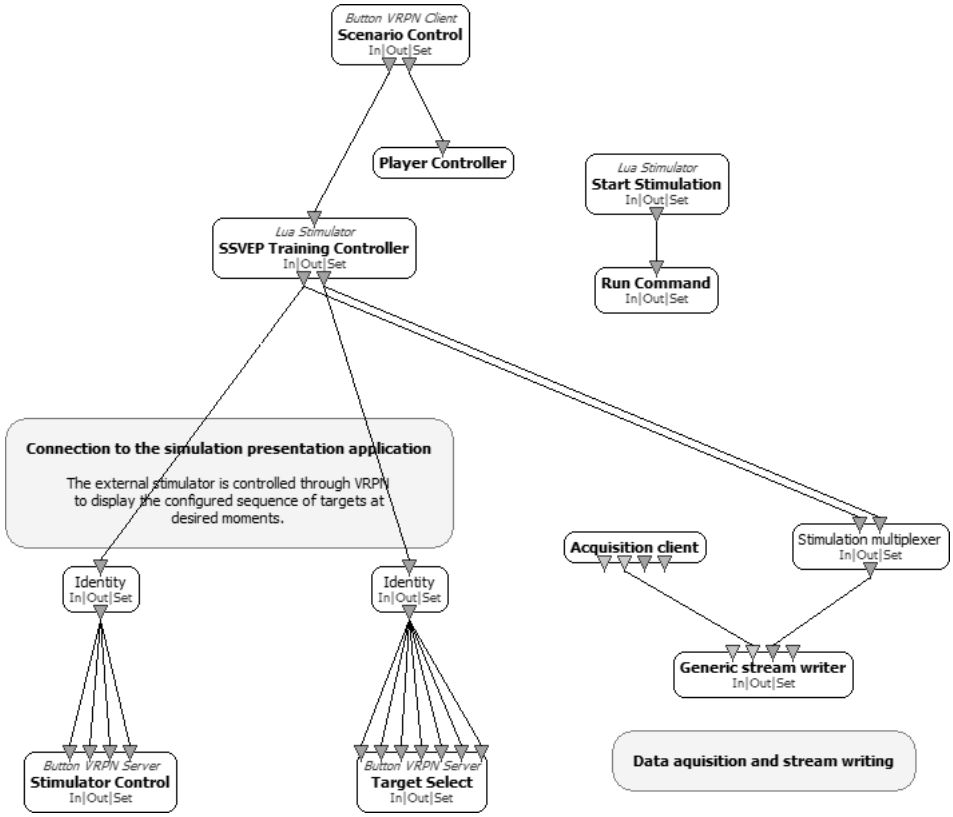


Figure 4.25: Aquisition of traning data

between target and non target states, and can then effectively differentiate these classes in an online environment. Classifier training is performed by the scenario, shown in Figure 4.27.

Since both the SVM and sLDA are binary classifiers, while dealing with the three class problem in this case, they are trained with a "one-versus all" paradigm i.e., the first classifier takes features from the first frequency stimulation as the target class, and stimulations from the other two frequencies as non-target. The same is true for the second and third classifier. A voting algorithm is then used to select the class from the three classifier outputs. If all the classifier output a non-target condition, then the state is said to be NC.

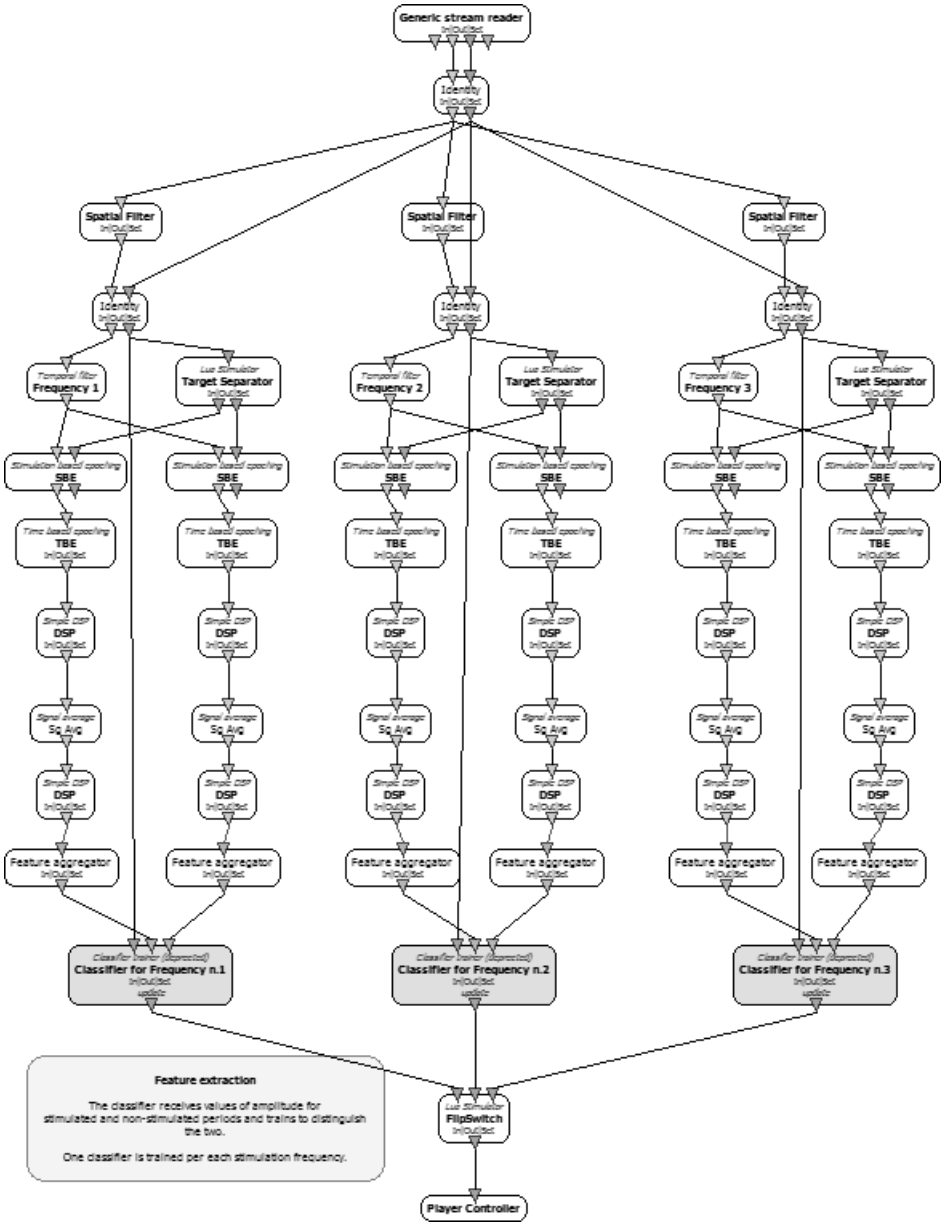


Figure 4.26: Classifier training scenario, using adaptive CSP filters and band power as features

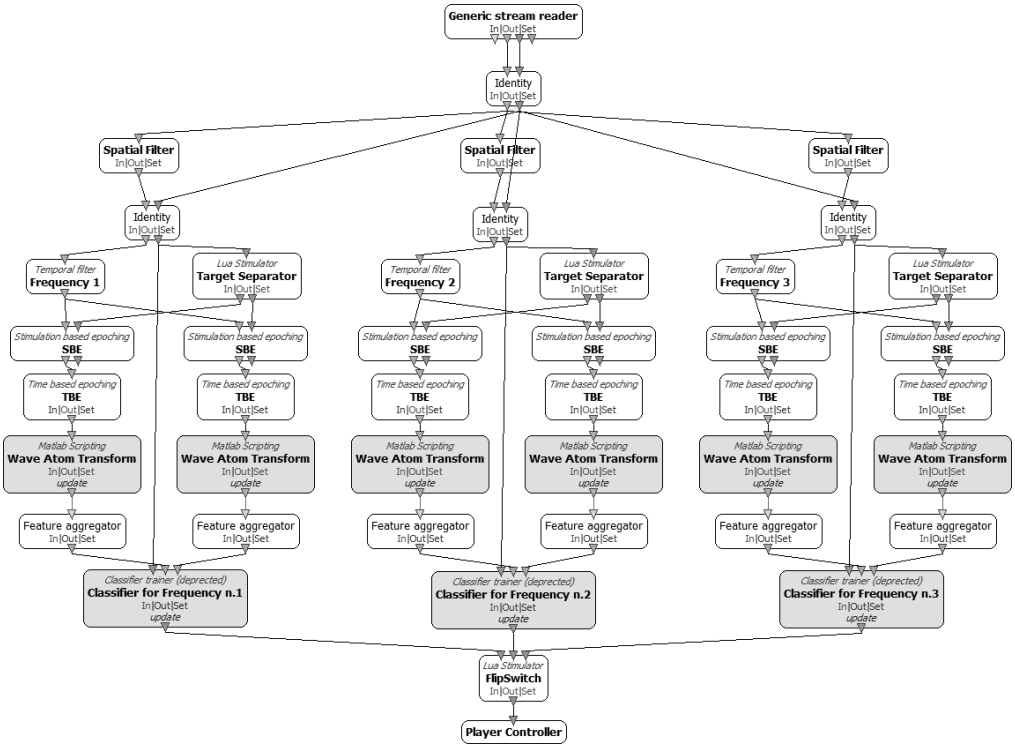


Figure 4.27: Classifier training, using WAT transform coefficients as features

4.7.4. Results

To evaluate the system, two datasets have been acquired from two different subjects, marked S1 and S2. Classification has been performed using two feature extraction methods for evaluation - the Wave Atom transform (WAT), and band power (BP) features.

Since there is a lack of training data, a 10 fold cross-validation is performed and accuracy is measured on the same data used for classifier training. The accuracy metric is chosen for the representation of results, since it is a simple metric that is directly linked to system usability by the user. It likely overestimates the classification result, since the classifier has been trained on the same data. The results most probably indicate higher performances than what the user will actually have during the online classification.

The results of classifier validation are presented in Table 7

Table 7: Comparison of classification accuracy

Classifier	Features	Accuracy, %		F	
		S1	S2	S1	S2
LDA	WAT	71.5	78.2	0.64	0.67
	BPS	66.2	73.2	0.56	0.62
sLDA	WAT	70.6	77.4	0.64	0.68
	BP	68.4	73.5	0.59	0.61
SVM, linear kernel	WAT	75.5	79.3	0.64	0.68
	BP	74.3	75.1	0.64	0.66
SVM, RBF kernel	WAT	**78.7**	**82.2**	0.68	0.71
	BP	74.0	77.4	0.63	0.67

These results indicate that the Wave Atom transform-based feature extraction method performed better than band power features on all 4 datasets. This method can also be used in the SSVEP paradigm.

Although the best results were achieved by using the SVM classifier with a linear kernel, results obtained with other classifiers are very similar. This shows that the choice of a good feature extraction algorithm is more important in BCI applications. A non-optimal classifier can produce good results, because most models pick up on good feature data. With good features, one can use a simpler classifier that runs faster.

These results also show that it is possible to develop a BCI interface system based on low-cost acquisition devices, such as the Emotiv EPOC, which performs at a reasonable usability level.

4.7.5. Online application

After performing classifier training, subjects are invited to participate in a video-game-like experiment. During this game, the subjects are presented with an interface from Figure 4.28. The "spaceship", comprised of two "engines" – the two rectangles and a "cannon" – the triangle. The subject is able to rotate the spaceship by focusing his/her attention to one of the rectangular targets.

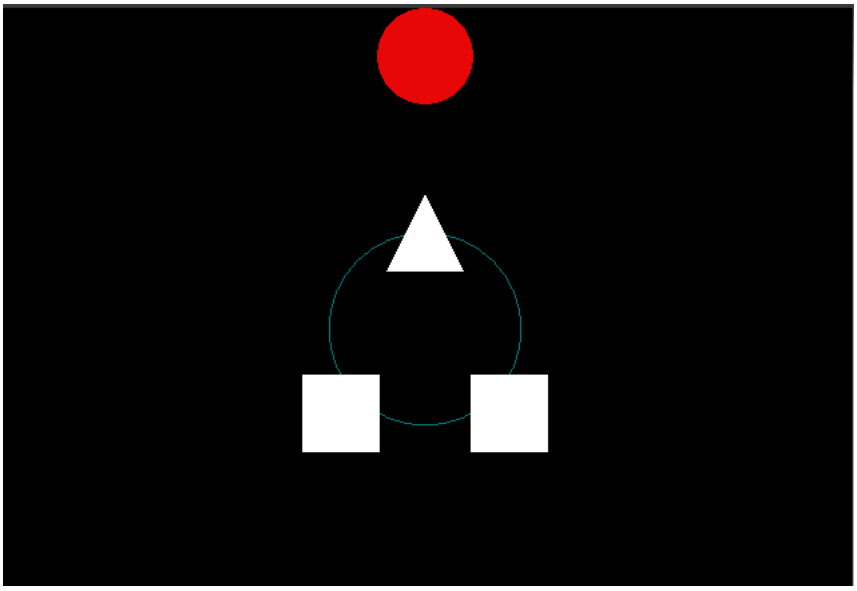


Figure 4.28: User interface of the online shooter game

Focusing on the left target makes the spaceship turn left, focusing on the right target – turn right. By focusing attention to the middle triangle, the user is able to fire the spaceship cannon. A red circular target appears next to the ship at a random location. The aim of the game is to rotate the spaceship and fire its canon to hit the red target. Once the target is hit, it disappears to reappear in another position.

The online game is executed using the OpenVibe scenario shown in Figure 4.29. The data is acquired in real time by the Acquisition client. The data is processed in the same way as the training data, in order to obtain the same feature vectors which the classifier can then identify. The output of the three classifiers is then input into a SSVEP voter algorithm, which decides on the class label of the current signal. If an NC state is detected, a class label of "0"

is assigned to the trial. The control signal can then be used to move the ship and is passed to the Ship Control application. While in NC state, no action is performed.

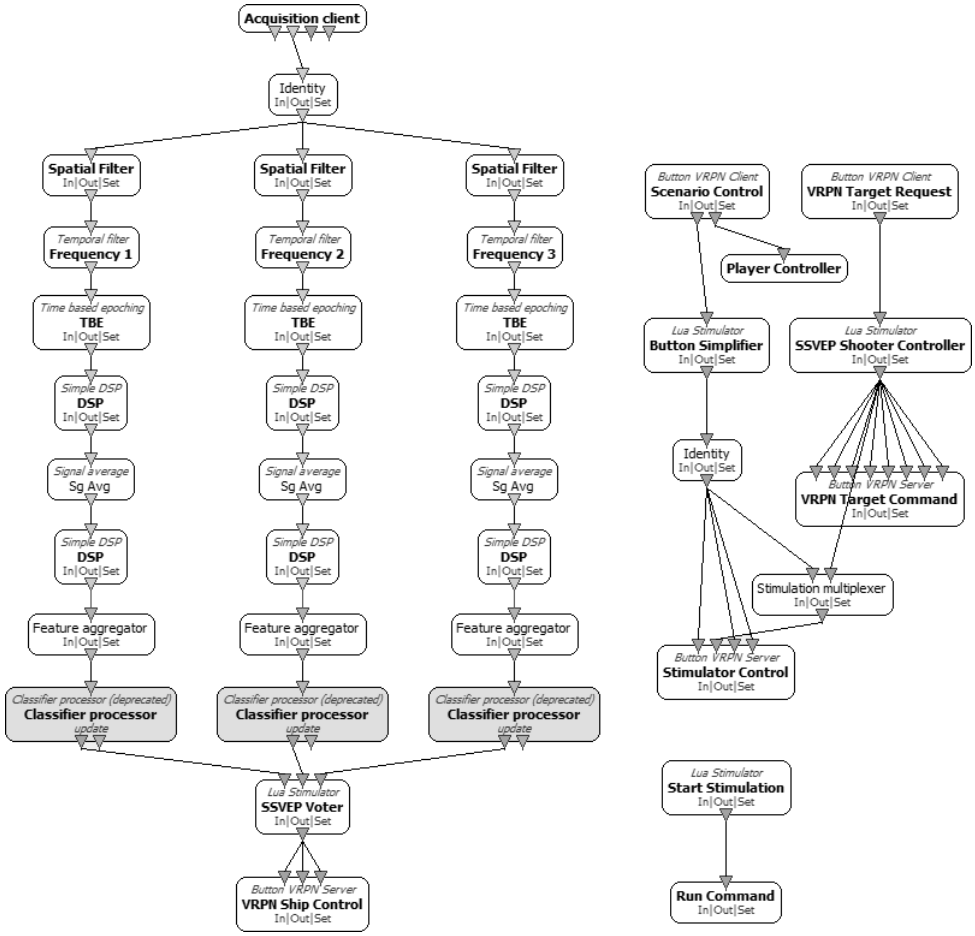


Figure 4.29: Online test shooter scenario

4.7.6. Future work

There are many challenges in the future of BCI. Currently none of the BCI systems are capable of helping disabled users with cursor control, which could help use ordinary computer applications. Accuracy, speed, usability and feedback methods should be improved in current BCI systems. Accuracy is the most important and greatly affects performance of the BCI, therefore feature extraction algorithms have to be developed to identify robust features, that do not get affected by user fatigue or a noisy environment.

Concerning the developed of our own BCI system, it would be interesting to let it be tested by a significant number of users, in order to analyze its generalization capabilities. Improving the EPOC headsets sensor placement could lead to a better results with MI-based tasks.

Concerning the signal processing and classification part of BCI design, we believe that a better approach would be to combine, rather than select preprocessing, feature extraction and classification methods. Numerous methods have been proposed and tested in the BCI domain, and while some of them have sometimes been proven to perform better than others, no single method has been identified as being the best. This is partly due to the differences in system users. Therefore, we should focus on combining existing methods together and adapting them to best suit the user. Since different methods exploit different aspects of EEG, these methods could be used together in a complementary way, and would probably lead to better results, than when using some "single best" method alone.

By improving system cost and ergonomics our results suggest that in the future, development of all-in-one headset solutions could be possible, which could effectively bring BCI to a higher level of applicability and usability. This also enables the use of BCI technology for the general public who can enjoy entertaining applications, games and virtual reality.

4.8. Conclusions

In this chapter, we have proposed several algorithms for EEG signal processing and classification.

By using the Wave Atom transform, we were able to achieve a 90 % accuracy, with a feature vector, comprised of only 4 features. This result is higher than the winner of the BBCI competition. This suggests that the method is efficient and retains the information needed for mental task classification. Also the small size of the feature vector allows for a small ANN architecture and complex training algorithms to be used. The best improvement in the classification was achieved using Bayesian regularization training. This improvement would mostly benefit real-time BCI applications, since it allows for fast classifier retraining, even when using large datasets. This algorithm has been chosen as the basis for the SSVEP based gaming system, described in detail in Chapter

4.7.

The Class-Adaptive denoising method uses Fisher (or Helliger) distance metric to evaluate the distance between shrinked frequency components in the time-frequency representations of signal data belonging to positive and negative dataset classes. To achieve maximal separation between positive and negative classes, the distance metric values are maximized using Nelder-Mead optimization. The optimized shrinkage function is used for EEG data denoising.

To show that this denoising procedure produces good classification results, the denoised data is then classified using an SVM classifier with a linear kernel.

Experimental results show that Class-Adaptive denoising can improve classification results as compared with the case were no signal denoising is used. The best denoising is achieved using three-parameter shrinkage functions (proposed by Yang, Wei and Atto) with their parameter values optimized for each frequency component of the frequency domain, while soft denoising has failed due to large bias of the denoised signal. The Fisher distance metric generally allows achieving better results than the Helliger distance metric. In fact, denoising using Helliger distance has failed to produce better results even when compared with the original (noisy) data.

In Section 4.5 we propose a novel nonlinear operator based on the generalization of the Teager-Kaiser Energy Operator, called Homogeneous Multivariate Polynomial Operator (HMPO). The applicability of the proposed operator is demonstrated for the classification of the EEG signals. The experimental results obtained using a SVM demonstrate an improvement of the classification results. The obtained results were comparable with those of BCI competition winners, showing that the HMPO operator can be effectively used for BCI feature extraction.

In Section 4.6 we have shown that efficient classification can be accomplished by time-bounding the ANN classifier and performing dataset reduction. Our results show this method to be effective for reducing data size without the loss of important signal information. Data classification was performed using artificial neural networks with various hidden layer sizes and training functions.

Since BCI classifiers require constant retraining in real-life applications, fast classifier training algorithms need to be developed. For this reason, we have proposed the Real-Time Voted Perceptron, a time-aware training algorithm for the Voted Perceptron. To cope with large amounts of the EEG data that BCI applications must cope with and to make fast training feasible, we also researched dataset transformations to reduce dataset sizes. Experimental results show that the size of data can be reduced significantly using signal down-sampling, without significant loss of classification quality, while satisfying the real-time constraints for NN training. Row sampling has a more significant negative effect on the classification quality, therefore, we do not suggest using it for BCI applications.

We have also proposed and implemented a BCI system, based on the Steady State Visually Evoked Potential (SSVEP) paradigm and the Emotiv EPOC neuroheadset device. Scenarios for classifier training, signal preprocessing and feature extraction. The Wave Atom transform, described in detail in section 4.4 was chosen for feature extraction.

We have created a scenario, enabling the user to control a virtual spaceship by thought. The scenario enables the user to issue 3 control commands and has a no-command (NC) state, allowing for self-paced control.

A evaluation of the system has been conducted using two subjects, that were new to any BCI system use. The first results suggest the efficiency of this method, as the results reached an accuracy over 80%. It is, of course, necessary to validate these results with more subjects, however the presented results are promising.

5. CONCLUSIONS

In this thesis, we have studied electroencephalogram (EEG) signal processing and classification techniques in order to design Brain-Computer Interface (BCI) systems to be used in a out-of-the-laboratory setting, with these main objectives:

1. improving efficiency in terms of accuracy of the BCI;
2. improving usability and applicability, therefore moving twoards the end-user;
3. designing a user friendly BCI system.

To reach these objectives, we have proposed several algorithms for EEG processing and classification.

1. Concerning signal preprocessing, we have proposed a Class-Adaptive signal denoising approach, improving classification results. 94.6% accuracy was achieved on a standard dataset. Because of complexity and the need for a lot of training data, this method was shown to be suited only for offline data analysis.
2. We have also proposed a Wave Atom transform algorithm for EEG feature extraction. The algorithm is shown to be effective in reducing the number of features, classifier training and testing time. This algorithm was tested on motor imagery (MI) data and also selected for a Steady-State Visually Evoked Potential (SSVEP) paradigm based system development. It was able to achieve an accuracy of 90% with a standard motor imagery (MI) dataset and showed good results (82.2% accuracy) with SSVEP data.
3. We have proposed a signal denoising method based on a Homogeneous Multivariate Polynomial Operator (HMPO). This method improved classification accuracy when used together with state-of-the-art classifiers, such as Support Vector Machines (SVM), to 82.8% on a standard dataset.
4. A algorithm for EEG signal fractional delay time embedding into a high dimensional phase space (see Chapter 4.3) has also been studied. The algorithm proposes a novel fractional time embedding scheme.
5. Finally, we have developed a 3 class BCI system, based on SSVEP and the Emotiv EPOC headset. An online target shooting game, implemented in the OpenViBE environment, has been used for feedback. The system utilizes Wave Atom transform for feature extraction. The system achieved an accuracy of 80.5% while using a Support Vector Machine classifier with a radial basis kernel.

Taken together, these results show that BCI can actually be used as an interaction technique for complex applications, providing real time operation and feedback.

These results also highlight that BCI can be feasible even when using a low resolution EEG acquisition devices. This allows for reduced system cost, mobility, subject preparation time and, consequently, allows for the subject to be prepared by a non-expert supervisor;

References

[Addison, 2005] Addison, P. S. (2005). Wavelet transforms and the ECG: a review. *Physiological Measurement*, 26(5):R155–R199.

[Aladjem, 1996] Aladjem, M. E. (1996). Two-class pattern discrimination via recursive optimization of patrick-fisher distance. In *Pattern Recognition, 1996., Proceedings of the 13th International Conference on*, volume 2, pages 60–64. IEEE.

[Allison and Neuper, 2010] Allison, B. and Neuper, C. (2010). Could anyone use a BCI? In *Brain-Computer Interfaces*, pages 35–54. Springer Science and Business Media.

[Álvarez Estévez et al., 2011] Álvarez Estévez, D., Sánchez-MaroñO, N., Alonso-Betanzos, A., and Moret-Bonillo, V. (2011). Reducing dimensionality in a database of sleep EEG arousals. *Expert Syst. Appl.*, 38(6):7746–7754.

[Anderson et al., 1998] Anderson, C., Stolz, E., and Shamsunder, S. (1998). Multivariate autoregressive models for classification of spontaneous electroencephalographic signals during mental tasks. *Biomedical Engineering, IEEE Transactions on*, 45(3):277–286.

[Ang et al., 2012] Ang, K. K., Guan, C., Phua, K. S., Wang, C., Teh, I., Chen, C. W., and Chew, E. (2012). Transcranial direct current stimulation and EEG-based motor imagery BCI for upper limb stroke rehabilitation. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pages 4128–4131.

[Athitsos et al., 2008] Athitsos, V., Papapetrou, P., Potamias, M., Kollios, G., and Gunopulos, D. (2008). Approximate embedding-based subsequence matching of time series. In *Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data*, SIGMOD '08, pages 365–378, New York, NY, USA. ACM.

[Atto et al., 2008] Atto, A., Pastor, D., and Mercier, G. (2008). Smooth sigmoid wavelet shrinkage for non-parametric estimation. In *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, pages 3265–3268.

[Bahari and Janghorbani, 2013] Bahari, F. and Janghorbani, A. (2013). EEG-based emotion recognition using recurrence plot analysis and k nearest neighbor classifier. In *Biomedical Engineering (ICBME), 2013 20th Iranian Conference on*, pages 228–233.

[Bai et al., 2014] Bai, X., Wang, X., Zheng, S., and Yu, M. (2014). The offline feature extraction of four-class motor imagery EEG based on ica and wavelet-csp. In *Control Conference (CCC), 2014 33rd Chinese*, pages 7189–7194.

[Bajaj and Pachori, 2014] Bajaj, V. and Pachori, R. (2014). Human emotion classification from EEG signals using multiwavelet transform. In *Medical Biometrics, 2014 International Conference on*, pages 125–130.

[Bar-Yossef et al., 2004] Bar-Yossef, Z., Jayram, T., Kumar, R., and Sivakumar, D. (2004). An information statistics approach to data stream and communication complexity. *Journal of Computer and System Sciences*, 68(4):702 – 732. Special Issue on {FOCS} 2002.

[Baranauskas, 2014] Baranauskas, G. (2014). What limits the performance of current invasive brain machine interfaces? *Frontiers in Systems Neuroscience*, 8(68).

[Barreto et al., 1996] Barreto, A., Taberner, A., and Vicente, L. (1996). Neural network classification of spatio-temporal EEG readiness potentials. In *Biomedical Engineering Conference, 1996., Proceedings of the 1996 Fifteenth Southern*, pages 73–76.

[Bashashati et al., 2007] Bashashati, A., Fatourechi, M., Ward, R. K., and Birch, G. (2007). A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals. *Journal of Neural engineering*, 4(2):35–57.

[Bayliss, 2001] Bayliss, J. D. (2001). *A Flexible Brain-computer Interface*. PhD thesis, University of Rochester. AAI3023727.

[Başar-Eroglu et al., 1996] Başar-Eroglu, C., Strüber, D., Kruse, P., Başar, E., and Stadler, M. (1996). Frontal gamma-band enhancement during multistable visual perception. *International Journal of Psychophysiology*, 24(1–2):113 – 125. New Advances in {EEG} and cognition.

[Ben-Ari, 1990] Ben-Ari, M. (1990). *Principles of Concurrent and Distributed Programming*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

[Bhattacharyya et al., 2011] Bhattacharyya, S., Khasnobish, A., Konar, A., Tibarewala, D., and Nagar, A. (2011). Performance analysis of left/right hand movement classification from EEG signal by intelligent algorithms. In *Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB), 2011 IEEE Symposium on*, pages 1–8.

[Bi et al., 2013] Bi, L., Jie, K., an Fan, X., and Li, Y. (2013). A SSVEP brain-computer interface with the hybrid stimuli of SSVEP and p300. In *Complex*

Medical Engineering (CME), 2013 ICME International Conference on, pages 211–214.

[Bifet et al., 2010] Bifet, A., Holmes, G., Pfahringer, B., Kranen, P., Kremer, H., Jansen, T., and Seidl, T. (2010). Moa: Massive online analysis, a framework for stream classification and clustering. In *Journal of Machine Learning Research (JMLR) Workshop and Conference Proceedings, Volume 11: Workshop on Applications of Pattern Analysis*, pages 44–50. Journal of Machine Learning Research.

[Birbaumer et al., 1999] Birbaumer, N., Flor, H., Ghanayim, N., Hinterberger, T., Iverson, I., Taub, E., Kotchoubey, B., Kübler, A., and Perelmouter, J. (1999). A brain-controlled spelling device for the completely paralyzed. *Nature*, (398):297–298.

[Birvinskas et al., 2012] Birvinskas, D., Jusas, V., Martisius, I., and Damasevicius, R. (2012). EEG dataset reduction and feature extraction using discrete cosine transform. In *Computer Modeling and Simulation (EMS), 2012 Sixth UKSim/AMSS European Symposium on*, pages 199–204.

[Birvinskas et al., 2013] Birvinskas, D., Jusas, V., Martišius, I., and Damaševičius, R. (2013). Data compression of EEG signals for artificial neural network classification. *Information technology and control*, 42(3):238–241.

[Blanchard and Blankertz, 2004] Blanchard, G. and Blankertz, B. (2004). BCI competition 2003-data set iia: spatial patterns of self-controlled brain rhythm modulations. *Biomedical Engineering, IEEE Transactions on*, 51(6):1062–1066.

[Blankertz et al., 2006a] Blankertz, B., Dornhege, G., Krauledat, M., Muller, K., Kunzmann, V., Losch, F., and Curio, G. (2006a). The berlin brain-computer interface: EEG-based communication without subject training. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 14(2):147–152.

[Blankertz et al., 2006b] Blankertz, B., Dornhege, G., Krauledat, M., Schröder, M., Williamson, J., Murray-Smith, R., and Müller, K.-R. (2006b). The berlin brain-computer interface presents the novel mental typewriter hex-o-spell.

[Blankertz et al., 2009] Blankertz, B., Sanelli, C., Halder, S., Hammer, E.-M., Kübler, A., Müller, K.-R., Curio, G., and Dickhaus, T. (2009). Predicting BCI performance to study BCI illiteracy. In *7th NFSI & ICBEM*.

[Blankertz et al., 2010] Blankertz, B., Tangermann, M., Vidaurre, C., Fazli, S., Sannelli, C., Haufe, S., Maeder, C., Ramsey, L. E., Sturm, I., Curio, G., and Mueller, K. R. (2010). The berlin brain-computer interface: Non-medical uses of BCI technology. *Frontiers in Neuroscience*, 4(198).

[Boashash et al., 2012] Boashash, B., Boubchir, L., and Azemi, G. (2012). Improving the classification of newborn EEG time-frequency representations using a combined time-frequency signal and image approach. In *Information Science, Signal Processing and their Applications (ISSPA), 2012 11th International Conference on*, pages 280–285.

[Boostani et al., 2007] Boostani, R., Graimann, B., Moradi, M., and Pfurtscheller, G. (2007). A comparison approach toward finding the best feature and classifier in cue-based BCI. *Medical & biological engineering & computing*, 45(4):403–412.

[Bos et al., 2010] Bos, D.-O., Reuderink, B., van de Laar, B., Gurkok, H., Muhl, C., Poel, M., Heylen, D., and Nijholt, A. (2010). Human-computer interaction for BCI games: Usability and user experience. In *Cyberworlds (CW), 2010 International Conference on*, pages 277–281. IEEE.

[Bostanov, 2004] Bostanov, V. (2004). BCI competition 2003-data sets ib and iib: feature extraction from event-related brain potentials with the continuous wavelet transform and the t-value scalogram. *Biomedical Engineering, IEEE Transactions on*, 51(6):1057–1061.

[Bruckhaus, 2007] Bruckhaus, T. (2007). The business impact of predictive analytics. In *Knowledge Discovery and Data Mining*, pages 114–138. {IGI} Global.

[Buzsaki, 2006] Buzsaki, G. (2006). *Rhythms of the Brain*. Oxford University Press.

[Caminhas et al., 2003] Caminhas, W. M., Vieira, D. A., and Vasconcelos, J. A. (2003). Parallel layer perceptron. *Neurocomputing*, 55(3–4):771 – 778. Evolving Solution with Neural Networks.

[Cao et al., 2012] Cao, T., Wan, F., Mak, P. U., Mak, P.-I., Vai, M. I., and Hu, Y. (2012). Flashing color on the performance of SSVEP-based brain-computer interfaces. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pages 1819–1822.

[Cappa and Vignolo, 1982] Cappa, S. F. and Vignolo, L. A. (1982). Locked-in syndrome for 12 years with preserved intelligence. *Annals of neurology*, 11(5):545–545.

[Carmena et al., 2003] Carmena, J. M., Lebedev, M. A., Crist, R. E., Doherty, J. E. O., Santucci, D. M., Dimitrov, D. F., Patil, P. G., Henriquez, C. S., and Nicolelis, M. A. L. (2003). Learning to control a brain - machine interface for reaching and grasping by primates. *Plos Biol*, 1(2):e2.

[Carrino et al., 2012] Carrino, F., Dumoulin, J., Mugellini, E., Khaled, O., and Ingold, R. (2012). A self-paced BCI system to control an electric wheelchair: Evaluation of a commercial, low-cost EEG device. In *Biosignals and Biorobotics Conference (BRC), 2012 ISSNIP*, pages 1–6.

[Chang et al., 2000] Chang, S. G., Yu, B., and Vetterli, M. (2000). Adaptive wavelet thresholding for image denoising and compression. *Image Processing, IEEE Transactions on*, 9(9):1532–1546.

[Chapin et al., 1999] Chapin, J. K., Moxon, K. A., Markowitz, R. S., and Nicolelis, M. A. L. (1999). Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex. *Nature Neuroscience*, 2(7):664–670.

[Chatlapalli et al., 2004] Chatlapalli, S., Nazeran, H., Melarkod, V., Krishnam, R., Estrada, E., Pamula, Y., and Cabrera, S. (2004). Accurate derivation of heart rate variability signal for detection of sleep disordered breathing in children. In *Engineering in Medicine and Biology Society, 2004. IEMBS '04. 26th Annual International Conference of the IEEE*, volume 1, pages 538–541.

[Chen et al., 2010] Chen, C., Song, W., Zhang, J., Hu, Z., and Xu, H. (2010). An adaptive feature extraction method for motor-imagery BCI systems. In *Computational Intelligence and Security (CIS), 2010 International Conference on*, pages 275–279.

[Chen et al., 2014] Chen, C.-Y., Wu, C.-W., Lin, C.-T., and Chen, S.-A. (2014). A novel classification method for motor imagery based on brain-computer interface. In *Neural Networks (IJCNN), 2014 International Joint Conference on*, pages 4099–4102.

[Cheung and Van Veen, 2008] Cheung, B. and Van Veen, B. (2008). Estimation of cortical multivariate autoregressive models for EEG/meg using an expectation-maximization algorithm. In *Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on*, pages 1235–1238.

[Chu et al., 2006] Chu, C. T., Kim, S. K., Lin, Y. A., Yu, Y., Bradski, G. R., Ng, A. Y., and Olukotun, K. (2006). Map-reduce for machine learning on multicore. In Schölkopf, B., Platt, J. C., and Hoffman, T., editors, *NIPS*, pages 281–288. MIT Press.

[Cohen, 1960] Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37–46.

[Congedo, 2006] Congedo, M. (2006). Subspace projection filters for real-time brain electromagnetic imaging. *Biomedical Engineering, IEEE Transactions on*, 53(8):1624–1634.

[Congedo et al., 2006] Congedo, M., Lotte, F., and Lécuyer, A. (2006). Classification of movement intention by spatially filtered electromagnetic inverse solutions. *Physics in Medicine and Biology*, 51(8):1971–1989.

[Cortes and Vapnik, 1995] Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3):273–297.

[Coyle et al., 2011] Coyle, D., Garcia, J., Satti, A., and McGinnity, T. (2011). EEG-based continuous control of a game using a 3 channel motor imagery BCI: BCI game. In *Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB), 2011 IEEE Symposium on*, pages 1–7.

[Damasevicius, 2008] Damasevicius, R. (2008). Splice site recognition in dna sequences using k-mer frequency based mapping for support vector machine with power series kernel. In *Complex, Intelligent and Software Intensive Systems, 2008. CISIS 2008. International Conference on*, pages 687–692.

[Damasevicius and Stuikys, 2009] Damasevicius, R. and Stuikys, V. (2009). Specification and generation of learning object sequences for e-learning using sequence feature diagrams and metaprogramming techniques. In *Advanced Learning Technologies, 2009. ICALT 2009. Ninth IEEE International Conference on*, pages 572–576.

[Damaševičius et al., 2014] Damaševičius, R., Martišius, I., Jusas, V., and Birvinskas, D. (2014). Fractional delay time embedding of EEG signals into high dimensional phase space. In *Electronics and Electrical Engineering*, volume 20.

[del R.Millan et al., 2000] del R.Millan, J., Mourino, J., Babiloni, F., Cincotti, F., Varsta, M., and Heikkonen, J. (2000). Local neural classifier for EEG-based recognition of mental tasks. In *Neural Networks, 2000. IJCNN 2000, Proceedings of the IEEE-INNS-ENNS International Joint Conference on*, volume 3, pages 632–636 vol.3.

[Demanet and Ying, 2007] Demanet, L. and Ying, L. (2007). Wave atoms and sparsity of oscillatory patterns. *Applied and Computational Harmonic Analysis*, 23(3):368 – 387.

[Dietrich et al., 2010] Dietrich, D., Lang, R., Bruckner, D., Fodor, G., and B., M. (2010). Limitations, possibilities and implications of brain-computer interfaces. In *Human System Interactions (HSI), 2010 3rd Conference on*, pages 722–726.

[Donoho, 1995] Donoho, D. L. (1995). De-noising by soft-thresholding. *IEEE Trans. Inf. Theor.*, 41(3):613–627.

[Donoho and Johnstone, 1995] Donoho, D. L. and Johnstone, I. M. (1995). Adapting to unknown smoothness via wavelet shrinkage. *Journal of the american statistical association*, 90(432):1200–1224.

[Dornhege, 2006] Dornhege, G. (2006). *Increasing Information Transfer Rates for Brain-Computer Interfacing*. PhD thesis, Mathematisch-Naturwissenschaftlichen Fakultat der Universitat Potsdam.

[Dornhege et al., 2004] Dornhege, G., Blankertz, B., Curio, G., and Müller, K.-R. (2004). Increase information transfer rates in BCI by csp extension to multi-class. *Advances in Neural Information Processing Systems*.

[Dua and Raj, 2012] Dua, G. and Raj, V. (2012). Mri denoising using waveatom shrinkage. *Global Journal of researches in engineering Electrical and electronics engineering*, 12(4).

[Duda et al., 2001] Duda, R. O., Hart, P. E., and Stork, D. G. (2001). *Pattern Classification*. Wiley, New York, 2 edition.

[Duvinage et al., 2012] Duvinage, M., Castermans, T., Dutoit, T., Petieau, M., Hoellinger, T., De Saedeleer, C., Seetharaman, K., and Cheron, G. (2012). A p300-based quantitative comparison between the emotiv epoc headset and a medical eeg device. *Biomedical Engineering*, 765:2012–764.

[Elbert et al., 1980] Elbert, T., Rockstroh, B., Lutzenberger, W., and Birbaumer, N. (1980). Biofeedback of slow cortical potentials. i. *Electroencephalography and Clinical Neurophysiology*, 48(3):293 – 301.

[Emotiv, 2014] Emotiv (2014). Emotiv epoc neuroheadset description. http://www.emotiv.com/epoc.php. Accessed: 2015.02.15.

[Erfanian and Gerivany, 2001] Erfanian, A. and Gerivany, M. (2001). EEG signals can be used to detect the voluntary hand movements by using an enhanced resource-allocating neural network. In *Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE*, volume 1, pages 721–724 vol.1.

[Falk et al., 2011] Falk, T., Guirgis, M., Power, S., and Chau, T. (2011). Taking nirs-BCIs outside the lab: Towards achieving robustness against environment noise. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 19(2):136–146.

[Faller et al., 2010] Faller, J., Müller-Putz, G., Schmalstieg, D., and Pfurtscheller, G. (2010). An application framework for controlling an avatar in a desktop-based virtual environment via a software SSVEP brain-computer interface. *Presence: Teleoperators and Virtual Environments*, 19(1):25–34.

[Farwell and Donchin, 1988] Farwell, L. and Donchin, E. (1988). Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology*, 70(6):510–523.

[Fatourechi et al., 2005] Fatourechi, M., Bashashati, A., Ward, R., and Birch, G. (2005). A hybrid genetic algorithm approach for improving the performance of the lf-asd brain computer interface. In *Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05). IEEE International Conference on*, volume 5, pages v/345–v/348 Vol. 5.

[Feng et al., 2013] Feng, M., Wang, X., and Zheng, S. (2013). A novel feature extraction method for motor imagery based on common spatial patterns with autoregressive parameters. In *Intelligent Control and Information Processing (ICICIP), 2013 Fourth International Conference on*, pages 225–230.

[Finke et al., 2009] Finke, A., Lenhardt, A., and Ritter, H. (2009). The mindgame: A p300-based brain–computer interface game. *Neural Networks*, 22(9):1329 – 1333. Brain-Machine Interface.

[Fisch and Spehlmann, 1999] Fisch, B. J. and Spehlmann, R. (1999). *Fisch and Spehlmann's EEG Primer: Basic Principles of Digital and Analog EEG*. Elsevier, 3 edition.

[Fisher, 1936] Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of eugenics*, 7(2):179–188.

[Freund and Shapire, 1999] Freund, Y. and Shapire, R. E. (1999). Large margin classification using the perceptron algorithm. *Machine Learning*, 37(3):277–296.

[Fu et al., 2011] Fu, Y., Xu, B., Pei, L., and Li, H. (2011). Reactive rhythm activities and offline classification of imagined speeds of finger movements. In *Bioinformatics and Biomedical Engineering, (iCBBE) 2011 5th International Conference on*, pages 1–5.

[Gama and Rodrigues, 2009] Gama, J. and Rodrigues, P. (2009). An overview on mining data streams. In Abraham, A., Hassanien, A.-E., de Leon F. de Carvalho, A., and Snášel, V., editors, *Foundations of Computational, IntelligenceVolume 6*, volume 206 of *Studies in Computational Intelligence*, pages 29–45. Springer Berlin Heidelberg.

[Geethanjali et al., 2012] Geethanjali, P., Mohan, Y., and Sen, J. (2012). Time domain feature extraction and classification of EEG data for brain computer interface. In *Fuzzy Systems and Knowledge Discovery (FSKD), 2012 9th International Conference on*, pages 1136–1139.

[Giannakakis and Nikita, 2008] Giannakakis, G. and Nikita, K. S. (2008). Estimation of time-varying causal connectivity on EEG signals with the use of adaptive autoregressive parameters. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 3512–3515.

[Goldberg, 2002] Goldberg, E. (2002). *The executive brain: Frontal lobes and the civilized mind*. Oxford University Press.

[Guger et al., 2015] Guger, C., Müller-Putz, G., and Allison, B. (2015). *Brain-Computer Interface Research: A State-of-the-Art Summary 4*. SpringerBriefs in Electrical and Computer Engineering. Springer International Publishing.

[Gulcar et al., 1998] Gulcar, H., Yilmaz, Y., and Demiralp, T. (1998). Classification of p300 component in single trial event related potentials. In *Biomedical Engineering Days, 1998. Proceedings of the 1998 2nd International Conference*, pages 48–50.

[Gürkök, 2012] Gürkök, H. (2012). *Mind the sheep! User experience evaluation & brain-computer interface games*. University of Twente.

[Hall et al., 2009] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H. (2009). The weka data mining software: An update. *SIGKDD Explor. Newsl.*, 11(1):10–18.

[Hamedi et al., 2014] Hamedi, M., Salleh, S.-H., Noor, A. M., and Mohammad-Rezazadeh, I. (2014). Neural network-based three-class motor imagery classification using time-domain features for BCI applications. In *Region 10 Symposium, 2014 IEEE*, pages 204–207.

[He and Musha, 1988] He, B. and Musha, T. (1988). Effects of cavity on EEG inverse dipole solution. In *Engineering in Medicine and Biology Society, 1988. Proceedings of the Annual International Conference of the IEEE*, pages 968–969 vol.2.

[Hjorth, 1970] Hjorth, B. (1970). {EEG} analysis based on time domain properties. *Electroencephalography and Clinical Neurophysiology*, 29(3):306 – 310.

[Hoffmann et al., 2005] Hoffmann, U., Garcia, G., Vesin, J., Diserens, K., and Ebrahimi, T. (2005). A boosting approach to p300 detection with application to brain-computer interfaces. In *Neural Engineering, 2005. Conference Proceedings. 2nd International IEEE EMBS Conference on*, pages 97–100.

[Holz et al., 2013] Holz, A. M., Botrel, L., and Kubler, A. (2013). Bridging gaps: Long-term independent BCI home- use by a locked-in end-user. *Proceedings of TOBI Workshop IV*.

[Hortal et al., 2013] Hortal, E., Ubeda, A., Ianez, E., Planelles, D., and Azorin, J. (2013). Online classification of two mental tasks using a svm-based BCI system. In *Neural Engineering (NER), 2013 6th International IEEE/EMBS Conference on*, pages 1307–1310.

[Hu et al., 2006] Hu, M., Li, J., Li, G., Tang, X., and Ding, Q. (2006). Classification of normal and hypoxia EEG based on approximate entropy and welch power-spectral-density. In *Neural Networks, 2006. IJCNN '06. International Joint Conference on*, pages 3218–3222.

[Hu et al., 2010] Hu, X.-S., Hong, K.-S., and Ge, S. (2010). Nirs based brain activation mapping: A real-time application. In *Control Automation and Systems (ICCAS), 2010 International Conference on*, pages 2254–2256.

[Huang et al., 2012] Huang, D., Qian, K., Fei, D.-Y., Jia, W., Chen, X., and Bai, O. (2012). Electroencephalography (EEG)-based brain computer interface (BCI): A 2-d virtual wheelchair control based on event-related desynchronization/synchronization and state control. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 20(3):379–388.

[Itoh and Shishido, 2008] Itoh, M. and Shishido, Y. (2008). Fisher information metric and poisson kernels. *Differential geometry and its applications*, 26(4):347–356.

[Jamoos et al., 2011] Jamoos, A., Grivel, E., Shakarneh, N., and Abdel-Nour, H. (2011). Dual optimal filters for parameter estimation of a multivariate autoregressive process from noisy observations. *Signal Processing, IET*, 5(5):471–479.

[Jasper, 1958] Jasper, H. H. (1958). Report of the committee on methods of clinical examination in electroencephalography. *Electroencephalography and Clinical Neurophysiology*, 10(2):370–375.

[Joachims, 2005] Joachims, T. (2005). A support vector method for multivariate performance measures. In *Proceedings of the 22nd International Conference on Machine Learning*, pages 377–384. ACM Press.

[Kaiser, 1990] Kaiser, J. (1990). On a simple algorithm to calculate the 'energy' of a signal. In *Acoustics, Speech, and Signal Processing, 1990. ICASSP-90., 1990 International Conference on*, pages 381–384 vol.1.

[Kapeller et al., 2012] Kapeller, C., Hintermüller, C., and Guger, C. (2012). Augmented control of an avatar using an SSVEP based BCI. In *Proceedings of the 3rd Augmented Human International Conference*, AH '12, pages 27:1–27:2, New York, NY, USA. ACM.

[Kaper et al., 2004] Kaper, M., Meinicke, P., Grossekathoefer, U., Lingner, T., and Ritter, H. (2004). BCI competition 2003-data set iib: support vector machines for the p300 speller paradigm. *Biomedical Engineering, IEEE Transactions on*, 51(6):1073–1076.

[Kashtiban et al., 2011] Kashtiban, A., Razmi, H., and Kozehkonan, M. (2011). Combined lvq neural network and multivariate statistical method employing wavelet coefficient for EEG signal classification. In *Mechatronics (ICM), 2011 IEEE International Conference on*, pages 809–814.

[Kennedy et al., 2000] Kennedy, P., Bakay, R., Moore, M., Adams, K., and Goldwaithe, J. (2000). Direct control of a computer from the human central nervous system. *Rehabilitation Engineering, IEEE Transactions on*, 8(2):198–202.

[Kennedy et al., 2004] Kennedy, P., Kirby, M., Moore, M., King, B., and Mallory, A. (2004). Computer control using human intracortical local field potentials. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 12(3):339–344.

[Khan et al., 2009] Khan, O., Kim, S.-H., Rasheed, T., Khan, A., and Kim, T.-S. (2009). Extraction of p300 using constrained independent component analysis. In *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pages 4031–4034.

[Klonowski et al., 1999] Klonowski, W., Jernajczyk, W., Niedzielska, K., RYDZ, A., and Stepien, R. (1999). Quantitative measure of complexity of EEG signal dynamics. *Acta Neurobiol Exp (Wars)*.

[Koo et al., 2014] Koo, B., Nam, Y., and Choi, S. (2014). A hybrid eog-p300 BCI with dual monitors. In *Brain-Computer Interface (BCI), 2014 International Winter Workshop on*, pages 1–4.

[Krishnam et al., 2005] Krishnam, R., Nazeran, H., Chatlapalli, S., Haltiwanger, E., and Pamula, Y. (2005). Detrended fluctuation analysis: A suitable long-term measure of hrv signals in children with sleep disordered breathing. In *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, pages 1174–1177.

[Kristo et al., 2015] Kristo, G., Höhne, J., Ortner, R., Reuderink, B., and Ramsey, N. (2015). Bnci horizon 2020.

[Krusienski et al., 2006] Krusienski, D., McFarland, D., and Wolpaw, J. (2006). An evaluation of autoregressive spectral estimation model order for brain-computer interface applications. In *Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE*, pages 1323–1326.

[Krusienski et al., 2011] Krusienski, D. J., Grosse-Wentrup, M., Galán, F., Coyle, D., Miller, K. J., Forney, E., and Anderson, C. W. (2011). Critical issues in state-of-the-art brain–computer interface signal processing. *Journal of neural engineering*, 8(2):025002.

[Kübler et al., 2001] Kübler, A., Kotchoubey, B., Kaiser, J., Wolpaw, J. R., and Birbaumer, N. (2001). Brain–computer communication: Unlocking the locked in. *Psychological bulletin*, 127(3):358.

[Kubler et al., 2006] Kubler, A., Mushahwar, V., Hochberg, L., and Donoghue, J. (2006). BCI meeting 2005-workshop on clinical issues and applications. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 14(2):131–134.

[Kvedalen, 2003] Kvedalen, E. (2003). Signal processing using the teager energy operator and other nonlinear operators. Master's thesis, University of Oslo.

[Lal et al., 2005] Lal, T. N., Hinterberger, T., Widman, G., Schröder, M., Hill, J., Rosenstiel, W., Elger, C. E., Schölkopf, B., and Birbaumer, N. (2005). Methods towards invasive human brain computer interfaces. In *IN ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 17*, pages 737–744. MIT Press.

[Lalor et al., 2005a] Lalor, E. C., Kelly, S. P., Finucane, C., Burke, R., Smith, R., Reilly, R. B., and McDarby, G. (2005a). Steady-state vep-based brain-computer interface control in an immersive 3d gaming environment. *EURASIP J. Appl. Signal Process.*, 2005:3156–3164.

[Lalor et al., 2005b] Lalor, E. C., Kelly, S. P., Finucane, C., Burke, R., Smith, R., Reilly, R. B., and McDarby, G. (2005b). Steady-state vep-based brain-computer interface control in an immersive 3d gaming environment. *EURASIP J. Appl. Signal Process.*, 2005:3156–3164.

[Lan et al., 2010] Lan, T., Erdogmus, D., Black, L., and van Santen, J. (2010). A comparison of different dimensionality reduction and feature selection methods for single trial erp detection. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pages 6329–6332.

[Lebedev and Nicolelis, 2006] Lebedev, M. A. and Nicolelis, M. A. (2006). Brain–machine interfaces: past, present and future. *TRENDS in Neurosciences*, 29(9):536–546.

[Lee et al., 2007] Lee, C.-W., Chen, D.-Y., Wu, C.-W., and Chen, J.-H. (2007). Comparing the spatial and temporal reproducibility of brain activation using three fmri techniques: Bold, fair, and vaso. In *Noninvasive Functional*

Source Imaging of the Brain and Heart and the International Conference on Functional Biomedical Imaging, 2007. NFSI-ICFBI 2007. Joint Meeting of the 6th International Symposium on, pages 258–261.

[Lee et al., 2011] Lee, P.-L., Yeh, C.-L., Cheng, J.-S., Yang, C.-Y., and Lan, G.-Y. (2011). An SSVEP-based BCI using high duty-cycle visual flicker. *Biomedical Engineering, IEEE Transactions on*, 58(12):3350–3359.

[Li et al., 2005] Li, Y., Dong, G., Gao, X., Gao, S., Ge, M., and Yan, W. (2005). Single trial EEG classification during finger movement task by using hidden markov models. In *Neural Engineering, 2005. Conference Proceedings. 2nd International IEEE EMBS Conference on*, pages 625–628.

[Li and Guan, 2006] Li, Y. and Guan, C. (2006). An extended em algorithm for joint feature extraction and classification in brain-computer interfaces. *Neural Computation*, 18(11):2730–2761.

[Li et al., 2013] Li, Y., Pan, J., Wang, F., and Yu, Z. (2013). A hybrid BCI system combining p300 and SSVEP and its application to wheelchair control. *Biomedical Engineering, IEEE Transactions on*, 60(11):3156–3166.

[Lin et al., 2010] Lin, C.-T., Ko, L.-W., Chang, M.-H., Duann, J.-R., Chen, J.-Y., Su, T.-P., and Jung, T.-P. (2010). Review of wireless and wearable electroencephalogram systems and brain-computer interfaces – a mini-review. *Gerontology*, 56(1):112–119.

[Liu et al., 2012] Liu, Y., Jiang, X., Cao, T., Wan, F., Mak, P. U., Mak, P.-I., and Vai, M. I. (2012). Implementation of ssvep based bci with emotiv epoc. In *Virtual Environments Human-Computer Interfaces and Measurement Systems (VECIMS), 2012 IEEE International Conference on*, pages 34–37.

[Lotte, 2008] Lotte, F. (2008). *Study of Electroencephalographic Signal Processing and Classification Techniques towards the use of Brain-Computer Interfaces in Virtual Reality Applications*. PhD thesis, l'Institut National des Sciences Appliquées de Rennes.

[Lotte et al., 2007] Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F., and Arnaldi, B. (2007). A review of classification algorithms for EEG-based brain-computer interfaces. *Journal of Neural Engineering*, 4(1):1–13.

[Lu et al., 2009] Lu, S., Guan, C., and Zhang, H. (2009). Unsupervised brain computer interface based on intersubject information and online adaptation. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 17(2):135–145.

[Lutzenberger et al., 1985] Lutzenberger, W., Elbert, T., Rockstroh, B., and Birbaumer, N. (1985). *The EEG*. Heidelberg, New York: Springer-Verlag.

[Maragos and Potamianos, 1995] Maragos, P. and Potamianos, A. (1995). Higher order differential energy operators. *Signal Processing Letters, IEEE*, 2(8):152–154.

[Martišius et al., 2013a] Martišius, I., Birvinskas, D., Damaševičius, R., and Jusas, V. (2013a). EEG dataset reduction and classification using wave atom transform. In Mladenov, V., Koprinkova-Hristova, P., Palm, G., Villa, A., Appollini, B., and Kasabov, N., editors, *Artificial Neural Networks and Machine Learning – ICANN 2013*, volume 8131 of *Lecture Notes in Computer Science*, pages 208–215. Springer Berlin Heidelberg.

[Martišius et al., 2011] Martišius, I., Birvinskas, D., Jusas, V., and Tamoševičius, v. (2011). A 2-d DCT hardware codec based on loeffler algorithm. *Electronics and Electrical Engineering*, 7(117):47–50.

[Martišius and Damaševičius, 2012] Martišius, I. and Damaševičius, R. (2012). Class-adaptive denoising for EEG data classification. In *Proceedings of the 11th International Conference on Artificial Intelligence and Soft Computing - Volume Part II*, ICAISC'12, pages 302–309, Berlin, Heidelberg. Springer-Verlag.

[Martišius et al., 2012] Martišius, I., Damaševičius, R., Jusas, V., and Birvinskas, D. (2012). Using higher order nonlinear operators for SVM classification of EEG data. *ELECTRONICS AND ELECTRICAL ENGINEERING*, 119(3):99–102.

[Martišius et al., 2013b] Martišius, I., šidlauskas, K., and Damaševičius, R. (2013b). Real-time training of voted perceptron for classification of EEG data. *International Journal of Artificial Intelligence (IJAI)*, 10(S13):41–50.

[Mason and Birch, 2003] Mason, S. and Birch, G. (2003). A general framework for brain-computer interface design. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 11(1):70–85.

[McCormick et al., 2010] McCormick, M., Ma, R., and Coleman, T. (2010). An analytic spatial filter and a hidden markov model for enhanced information transfer rate in EEG-based brain computer interfaces. In *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pages 602–605.

[McDonald et al., 2010] McDonald, R., Hall, K., and Mann, G. (2010). Distributed training strategies for the structured perceptron. In *Human Language Technologies: The 2010 Annual Conference of the North American*

Chapter of the Association for Computational Linguistics, HLT '10, pages 456–464, Stroudsburg, PA, USA. Association for Computational Linguistics.

[McFarland et al., 2006] McFarland, D., Anderson, C., Muller, K.-R., Schlogl, A., and Krusienski, D. (2006). BCI meeting 2005-workshop on BCI signal processing: feature extraction and translation. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 14(2):135–138.

[McFarland et al., 1997] McFarland, D. J., McCane, L. M., David, S. V., and Wolpaw, J. R. (1997). Spatial filter selection for EEG-based communication. *Electroencephalography and Clinical Neurophysiology*, 103(3):386 – 394.

[Menon et al., 2013] Menon, S., Brantner, G., Aholt, C., Kay, K., and Khatib, O. (2013). Haptic fmri: Combining functional neuroimaging with haptics for studying the brain's motor control representation. In *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, pages 4137–4142.

[Mensh et al., 2004] Mensh, B., Werfel, J., and Seung, H. (2004). BCI competition 2003-data set ia: combining gamma-band power with slow cortical potentials to improve single-trial classification of electroencephalographic signals. *Biomedical Engineering, IEEE Transactions on*, 51(6):1052–1056.

[Michahial et al., 2012] Michahial, S., Ranjith Kumar, R., Hemath Kumar, P., and Puneeth Kumar, A. (2012). Hand rotate EEG signal feature extraction by second order daubechies wavelet transform (dwt). In *Computing Communication Networking Technologies (ICCCNT), 2012 Third International Conference on*, pages 1–6.

[Middendorf et al., 2000] Middendorf, M., McMillan, G., Calhoun, G., and Jones, K. (2000). Brain-computer interfaces based on the steady-state visual-evoked response. *Rehabilitation Engineering, IEEE Transactions on*, 8(2):211–214.

[Mohammed et al., 2010] Mohammed, A., Jonathan Wu, Q., and Sid-Ahmed, M. (2010). Application of wave atoms decomposition and extreme learning machine for fingerprint classification. In Campilho, A. and Kamel, M., editors, *Image Analysis and Recognition*, volume 6112 of *Lecture Notes in Computer Science*, pages 246–255. Springer Berlin Heidelberg.

[Møller, 1993] Møller, M. F. (1993). A scaled conjugate gradient algorithm for fast supervised learning. *Neural networks*, 6(4):525–533.

[Moore et al., 1996] Moore, M., Mitra, S., and Bernstein, R. (1996). A generalization of the teager algorithm. *IEEE Transactions for Image Processing*, 5(6):950–963.

[Mrázek et al., 2005] Mrázek, P., Weickert, J., and Steidl, G. (2005). Diffusion-inspired shrinkage functions and stability results for wavelet denoising. *International Journal of Computer Vision*, 64(2-3):171–186.

[Mu et al., 2009] Mu, Z., Xiao, D., and Hu, J. (2009). Classification of motor imagery EEG signals based on stfts. In *Image and Signal Processing, 2009. CISP'09. 2nd International Congress on*, pages 1–4. IEEE.

[Mugler et al., 2010] Mugler, E., Ruf, C., Halder, S., Bensch, M., and Kubler, A. (2010). Design and implementation of a p300-based brain-computer interface for controlling an internet browser. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 18(6):599–609.

[Muller et al., 2001] Muller, K., Mika, S., Ratsch, G., Tsuda, K., and Scholkopf, B. (2001). An introduction to kernel-based learning algorithms. *IEEE Transactions on Neural Networks*, 12(2):181–201.

[Müller et al., 2004] Müller, K.-R., Krauledat, M., Dornhege, G., Curio, G., and Blankertz, B. (2004). Machine learning techniques for brain-computer interfaces. *Biomedical Technologies*, 49:11–22.

[Muller et al., 2011] Muller, S., Bastos-Filho, T., and Sarcinelli-Filho, M. (2011). Using a SSVEP-BCI to command a robotic wheelchair. In *Industrial Electronics (ISIE), 2011 IEEE International Symposium on*, pages 957–962.

[Muller-Putz et al., 2012] Muller-Putz, G., Klobassa, D., Pokorny, C., Pichler, G., Erlbeck, H., Real, R., Kubler, A., Risetti, M., and Mattia, D. (2012). The auditory p300-based SSBCI: A door to minimally conscious patients? In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pages 4672–4675.

[Nelder and Mead, 1965] Nelder, J. A. and Mead, R. (1965). A simplex method for function minimization. *The Computer Journal*, 7(4):308–313.

[NeuroSky, 2014] NeuroSky (2014). Neurosky mindwave neuroheadset description. http://neurosky.com/products-markets/{EEG}-biosensors/hardware/. Accessed: 2015.02.15.

[Nguyen et al., 2015a] Nguyen, T., Khosravi, A., Creighton, D., and Nahavandi, S. (2015a). {EEG} signal classification for {BCI} applications by wavelets and interval type-2 fuzzy logic systems. *Expert Systems with Applications*, 42(9):4370 – 4380.

[Nguyen et al., 2015b] Nguyen, T., Khosravi, A., Creighton, D., and Nahavandi, S. (2015b). Fuzzy system with tabu search learning for classification of motor imagery data. *Biomedical Signal Processing and Control*, 20:61 – 70.

[Nicolelis et al., 2000] Nicolelis, M. A. L., Wessberg, J., Stambaugh, C. R., Kralik, J. D., Beck, P. D., Laubach, M., Chapin, J. K., Kim, J., Biggs, S. J., and Srinivasan, M. A. (2000). Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. *Nature*, 408(6810):361–365.

[Norouzzadeh and Jampour, 2011] Norouzzadeh, Y. and Jampour, M. (2011). A novel curvelet thresholding function for additive gaussian noise removal. *International Journal of Computer Theory and Engineering*, 3(4).

[Pal et al., 2014] Pal, M., Khasnobish, A., Konar, A., Tibarewala, D., and Janarthanan, R. (2014). Performance enhancement of object shape classification by coupling tactile sensing with EEG. In *Electronics, Communication and Instrumentation (ICECI), 2014 International Conference on*, pages 1–4.

[Palaniappan, 2005] Palaniappan, R. (2005). Brain computer interface design using band powers extracted during mental tasks. In *Neural Engineering, 2005. Conference Proceedings. 2nd International IEEE EMBS Conference on*, pages 321–324.

[Palaniappan, 2006] Palaniappan, R. (2006). Utilizing gamma band to improve mental task based brain-computer interface design. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 14(3):299–303.

[Palaniappan and Raveendran, 2002] Palaniappan, R. and Raveendran, P. (2002). Genetic algorithm to select features for fuzzy artmap classification of evoked EEG. In *Circuits and Systems, 2002. APCCAS '02. 2002 Asia-Pacific Conference on*, volume 2, pages 53–56 vol.2.

[Palaniappan et al., 2000a] Palaniappan, R., Raveendran, P., Nishida, S., and Saiwaki, N. (2000a). Evolutionary fuzzy artmap for autoregressive model order selection and classification of EEG signals. In *Systems, Man, and Cybernetics, 2000 IEEE International Conference on*, volume 5, pages 3682–3686 vol.5.

[Palaniappan et al., 2000b] Palaniappan, R., Raveendran, P., Nishida, S., and Saiwaki, N. (2000b). Fuzzy atrmap classification of mental tasks using segmented and overlapped EEG signals. In *TENCON 2000. Proceedings*, volume 2, pages 388–391 vol.2.

[Park et al., 2013] Park, C., Looney, D., ur Rehman, N., Ahrabian, A., and Mandic, D. (2013). Classification of motor imagery BCI using multivariate empirical mode decomposition. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 21(1):10–22.

[Patwari et al., 2014] Patwari, A. K., Pansari, D., and Singh, V. P. (2014). Analysis of ECG signal compression technique using discrete wavelet transform for different wavelets. *International Journal of Engineering Trends and Technology*, 8(4):168–173.

[Peck and Van Ness, 1982] Peck, R. and Van Ness, J. (1982). The use of shrinkage estimators in linear discriminant analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (5):530–537.

[Penny and Roberts, 1999] Penny, W. and Roberts, S. (1999). EEG-based communication via dynamic neural network models. In *Neural Networks, 1999. IJCNN '99. International Joint Conference on*, volume 5, pages 3586–3590 vol.5.

[Pfurtscheller, 1992] Pfurtscheller, G. (1992). Event-related synchronization (ERS): an electrophysiological correlate of cortical areas at rest. *Electroencephalography and Clinical Neurophysiology*, 83:62–69.

[Pfurtscheller and Aranibar, 1977] Pfurtscheller, G. and Aranibar, A. (1977). Event-related cortical desynchronization detected by power measurements of scalp {EEG}. *Electroencephalography and Clinical Neurophysiology*, 42(6):817 – 826.

[Pfurtscheller et al., 1993] Pfurtscheller, G., Flotzinger, D., and Kalcher, J. (1993). Brain-computer interface—a new communication device for handicapped persons. *Journal of Microcomputer Applications*, 16(3):293 – 299.

[Pfurtscheller et al., 2005] Pfurtscheller, G., Müller-Putz, G. R., Pfurtscheller, J., and Rupp, R. (2005). EEG-based asynchronous BCI controls functional electrical stimulation in a tetraplegic patient. *EURASIP J. Appl. Signal Process.*, 2005:3152–3155.

[Pfurtscheller and Neuper, 2001] Pfurtscheller, G. and Neuper, C. (2001). Motor imagery and direct brain-computer communication. *Proceedings of the IEEE*, 89(7):1123–1134.

[Pfurtscheller et al., 1998] Pfurtscheller, G., Neuper, C., Schlogl, A., and Lugger, K. (1998). Separability of EEG signals recorded during right and left motor imagery using adaptive autoregressive parameters. *Rehabilitation Engineering, IEEE Transactions on*, 6(3):316–325.

[Pineda et al., 2003] Pineda, J., Silverman, D., Vankov, A., and Hestenes, J. (2003). Learning to control brain rhythms: making a brain-computer interface possible. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 11(2):181–184.

[Pinegger et al., 2015] Pinegger, A., Faller, J., Halder, S., Wriessnegger, S. C., and Müller-Putz, G. R. (2015). Control or non-control state: that is the question! an asynchronous visual p300-based BCI approach. *Journal of neural engineering*, 12(1):014001.

[Poornachandra and Kumaravel, 2008] Poornachandra, S. and Kumaravel, N. (2008). A novel method for the elimination of power line frequency in {ECG} signal using hyper shrinkage function. *Digital Signal Processing*, 18(2):116 – 126.

[Poornachandra and N., 2006] Poornachandra, S. and N., K. (2006). Subband-adaptive shrinkage for denoising of ecg signals. *EURASIP Journal on Applied Signal Processing*, 2006:1–9.

[Popescu et al., 2008] Popescu, F., Blankertz, B., and Müller, K.-R. (2008). Computational challenges for noninvasive brain computer interfaces. *IEEE Intelligent Systems*, 23(3):78–79.

[Punsawad and Wongsawat, 2013] Punsawad, Y. and Wongsawat, Y. (2013). Hybrid SSVEP-motion visual stimulus based BCI system for intelligent wheelchair. In *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, pages 7416–7419.

[Qassim et al., 2012] Qassim, Y., Cutmore, T., James, D., and Rowlands, D. (2012). Fpga implementation of morlet continuous wavelet transform for EEG analysis. In *Computer and Communication Engineering (ICCCE), 2012 International Conference on*, pages 59–64.

[Rafiee et al., 2009] Rafiee, J., Schoen, M., Prause, N., Urfer, A., and Rafiee, M. (2009). A comparison of forearm emg and psychophysical EEG signals using statistical signal processing. In *Computer, Control and Communication, 2009. IC4 2009. 2nd International Conference on*, pages 1–5.

[Rajeesh et al., 2010] Rajeesh, J., Moni, R. S., Kumar, S., and Gopalakrishnan, T. (2010). Rician noise removal on mri using wave atom transform with histogram based noise variance estimation. In *Communication Control and Computing Technologies (ICCCCT), 2010 IEEE International Conference on*, pages 531–535.

[Rakotomamonjy et al., 2005] Rakotomamonjy, A., Guigue, V., Mallet, G., and Alvarado, V. (2005). Ensemble of svms for improving brain computer interface p300 speller performances. In Duch, W., Kacprzyk, J., Oja, E., and Zadrożny, S., editors, *Artificial Neural Networks: Biological Inspirations – ICANN 2005*, volume 3696 of *Lecture Notes in Computer Science*, pages 45–50. Springer Berlin Heidelberg.

[Ramoser et al., 2000] Ramoser, H., Muller-Gerking, J., and Pfurtscheller, G. (2000). Optimal spatial filtering of single trial EEG during imagined hand movement. *Rehabilitation Engineering, IEEE Transactions on*, 8(4):441–446.

[Rance et al., 1995] Rance, G., Rickards, F. W., Cohen, L. T., De Vidi, S., and Clark, G. M. (1995). The automated prediction of hearing thresholds in sleeping subjects using auditory steady-state evoked potentials. *Ear and hearing*, 16(5):499–507.

[Rao and Derakhshani, 2005] Rao, R. and Derakhshani, R. (2005). A comparison of EEG preprocessing methods using time delay neural networks. In *Neural Engineering, 2005. Conference Proceedings. 2nd International IEEE EMBS Conference on*, pages 262–264.

[Raudonis et al., 2008] Raudonis, V., Narvydas, G., and Simutis, R. (2008). A classification of flash evoked potentials based on artificial neural network. *ELECTRONICS AND ELECTRICAL ENGINEERING*, 81(1):31–36.

[Raudys and Jain, 1991] Raudys, S. and Jain, A. (1991). Small sample size effects in statistical pattern recognition: recommendations for practitioners. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 13(3):252–264.

[Ray and Maunsell, 2011] Ray, S. and Maunsell, J. H. R. (2011). Different origins of gamma rhythm and high-gamma activity in macaque visual cortex. *PLoS Biol*, 9(4):e1000610.

[Regan, 1977] Regan, D. (1977). Steady-state evoked potentials. *JOSA*, 67(11):1475–1489.

[Reilly, 2013] Reilly, B. (2013). Mu rhythm. In Volkmar, F., editor, *Encyclopedia of Autism Spectrum Disorders*, pages 1940–1941. Springer New York.

[Renard et al., 2010] Renard, Y., Lotte, F., Gibert, G., Congedo, M., Maby, E., Delannoy, V., Bertrand, O., and Lécuyer, A. (2010). Openvibe: An open-source software platform to design, test, and use brain-computer interfaces in real and virtual environments. *Presence*, 19(1):35–53.

[Resalat et al., 2012] Resalat, S., Saba, V., Afdideh, F., and Heidarnejad, A. (2012). High-speed SSVEP-based BCI: Study of various frequency pairs and inter-sources distances. In *Biomedical and Health Informatics (BHI), 2012 IEEE-EMBS International Conference on*, pages 220–223.

[Reshmi and Amal, 2013] Reshmi, G. and Amal, A. (2013). Design of a BCI system for piloting a wheelchair using five class mi based EEG. In *Advances in Computing and Communications (ICACC), 2013 Third International Conference on*, pages 25–28.

[Rockstroh, 1989] Rockstroh, B. (1989). Area-specific regulation of slow cortical potentials. In Başar, E. and Bullock, T., editors, *Brain Dynamics*, volume 2 of *Springer Series in Brain Dynamics*, pages 467–477. Springer Berlin Heidelberg.

[Rohani et al., 2014] Rohani, D. A., Sorensen, H. B., and Puthusserypady, S. (2014). Brain-computer interface using p300 and virtual reality: A gaming approach for treating adhd. In *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, pages 3606–3609.

[Roy et al., 2013] Roy, R., Bonnet, S., Charbonnier, S., and Campagne, A. (2013). Mental fatigue and working memory load estimation: Interaction and implications for EEG-based passive BCI. In *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, pages 6607–6610.

[Samar et al., 1999] Samar, V. J., Bopardikar, A., Rao, R., and Swartz, K. (1999). Wavelet analysis of neuroelectric waveforms: A conceptual tutorial. *Brain and Language*, 66(1):7 – 60.

[Schloegl et al., 1997] Schloegl, A., Lugger, K., and Pfurtscheller, G. (1997). Using adaptive autoregressive parameters for a brain-computer-interface experiment. In *Engineering in Medicine and Biology Society, 1997. Proceedings of the 19th Annual International Conference of the IEEE*, volume 4, pages 1533–1535 vol.4.

[Shete and Shriram, 2014] Shete, S. and Shriram, R. (2014). Comparison of sub-band decomposition and reconstruction of EEG signal by daubechies9 and symlet9 wavelet. In *Communication Systems and Network Technologies (CSNT), 2014 Fourth International Conference on*, pages 856–861.

[Shyu et al., 2013] Shyu, K.-K., Chiu, Y.-J., Lee, P.-L., Liang, J.-M., and Peng, S.-H. (2013). Adaptive SSVEP-based BCI system with frequency and pulse duty-cycle stimuli tuning design. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 21(5):697–703.

[Singla et al., 2013] Singla, R., Khosla, A., and Jha, R. (2013). Influence of stimuli color on steady-state visual evoked potentials based BCI wheelchair control. *JBiSE*, 06(11):1050–1055.

[Smith, 1999] Smith, S. W. (1999). *The Scientist & Engineer's Guide to Digital Signal Processing*. California Technical Publishing.

[Song et al., 2012] Song, H., Zhang, D., Ling, Z., Zuo, H., and Hong, B. (2012). High gamma oscillations enhance the subdural visual speller. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pages 1711–1714.

[Song et al., 2013] Song, X., Yoon, S.-C., and Perera, V. (2013). Adaptive common spatial pattern for single-trial EEG classification in multisubject BCI. In *Neural Engineering (NER), 2013 6th International IEEE/EMBS Conference on*, pages 411–414.

[Souza et al., 2012] Souza, A., Filho, S., Felix, L., Maia, C., and Tierra-Criollo, C. (2012). Classification of imaginary movements using the magnitude-squared coherence feature extractor. In *Biosignals and Biorobotics Conference (BRC), 2012 ISSNIP*, pages 1–6.

[Subasi and Erçelebi, 2005] Subasi, A. and Erçelebi, E. (2005). Classification of {EEG} signals using neural network and logistic regression. *Computer Methods and Programs in Biomedicine*, 78(2):87 – 99.

[Suk and Lee, 2010] Suk, H.-I. and Lee, S.-W. (2010). Two-layer hidden markov models for multi-class motor imagery classification. In *Brain Decoding: Pattern Recognition Challenges in Neuroimaging (WBD), 2010 First Workshop on*, pages 5–8.

[Sulaiman et al., 2011] Sulaiman, N., Taib, M., Lias, S., Murat, Z., Aris, S., and Hamid, N. (2011). EEG-based stress features using spectral centroids technique and k-nearest neighbor classifier. In *Computer Modelling and Simulation (UKSim), 2011 UkSim 13th International Conference on*, pages 69–74.

[Sun et al., 2009] Sun, G., Li, K., Li, X., Zhang, B., Yuan, S., and Wu, G. (2009). A general framework of brain-computer interface with visualization and virtual reality feedback. In *Dependable, Autonomic and Secure Computing, 2009. DASC '09. Eighth IEEE International Conference on*, pages 418–423.

[Sutter, 1992] Sutter, E. E. (1992). The brain response interface: communication through visually-induced electrical brain responses. *Journal of Microcomputer Applications*, 15(1):31 – 45. Special Issue on Computers for Handicapped People.

[Takeda et al., 2008] Takeda, Y., Yamanaka, K., and Yamamoto, Y. (2008). Temporal decomposition of {EEG} during a simple reaction time task into stimulus- and response-locked components. *NeuroImage*, 39(2):742 – 754.

[Takens, 1981] Takens, F. (1981). Detecting strange attractors in turbulence. In Rand, D. and Young, L.-S., editors, *Dynamical Systems and Turbulence, Warwick 1980*, volume 898 of *Lecture Notes in Mathematics*, pages 366–381. Springer Berlin Heidelberg.

[Talukdar et al., 2014] Talukdar, M., Sakib, S., Pathan, N., and Fattah, S. (2014). Motor imagery EEG signal classification scheme based on autoregressive reflection coefficients. In *Informatics, Electronics Vision (ICIEV), 2014 International Conference on*, pages 1–4.

[Tanaka et al., 2005] Tanaka, K., Matsunaga, K., and Wang, H. (2005). Electroencephalogram-based control of an electric wheelchair. *Robotics, IEEE Transactions on*, 21(4):762–766.

[Teager, 1980] Teager, H. (1980). Some observations on oral air flow during phonation. *Acoustics, Speech and Signal Processing, IEEE Transactions on*.

[Teo et al., 2006] Teo, E., Huang, A., Lian, Y., Guan, C., Li, Y., and Zhang, H. (2006). Media communication center using brain computer interface. In *Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE*, pages 2954–2957.

[Thompson et al., 2013] Thompson, D. E., Blain-Moraes, S., and Huggins, J. E. (2013). Performance assessment in brain-computer interface-based augmentative and alternative communication. {*BioMed*} *Eng* {*OnLine*}, 12(1):43.

[Tomar and Patil, 2008] Tomar, V. and Patil, H. A. (2008). On the development of variable length teager energy operator (vteo). In *INTERSPEECH'08*, pages 1056–1059.

[Townsend et al., 2004] Townsend, G., Graimann, B., and Pfurtscheller, G. (2004). Continuous eeg classification during motor imagery-simulation of an asynchronous bci. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 12(2):258–265.

[Ulbikas et al., 1998] Ulbikas, J., Cenys, A., and Sulimova, O. (1998). Chaos parameters for EEG analysis. *Nonlinear Analysis. Modelling and Control*, 3:141–148.

[Unde and Shriram, 2014] Unde, S. and Shriram, R. (2014). Coherence analysis of EEG signal using power spectral density. In *Communication Systems and Network Technologies (CSNT), 2014 Fourth International Conference on*, pages 871–874.

[Ursulean and Lazar, 2007] Ursulean, R. and Lazar, A. M. (2007). Segmentation of electroencephalographic signals using an optimal orthogonal linear

prediction algorithm. *ELECTRONICS AND ELECTRICAL ENGINEERING*, 73(1):73–76.

[Vapnik, 1998] Vapnik, V. N. (1998). *Statistical learning theory*, volume 2. Wiley New York.

[Vasiljevas et al., 2012] Vasiljevas, M., Turčinas, R., and Martišius, I. (2012). Netiesinių operatorių taikymas EEG duomenų apdorojimui. In *Informacinės technologijos : 17-oji tarpuniversitetinė magistrantų ir doktorantų konferencija : konferencijos pranešimų medžiaga*, pages 23–26. Vilniaus universiteto Kauno humanitarinis fakultetas, Kauno technologijos universitetas, Vytauto Didžiojo universitetas, Vilniaus universitetas.

[Verikas and Gelžinis, 2008] Verikas, A. and Gelžinis, A. (2008). *Neuroniniai tinklai ir neuroniniai skaičiavimai : mokomoji knyga*. Technologija.

[Vidal, 1973] Vidal, J. J. (1973). Toward direct brain-computer communication. *Annual Review of Biophysics and Bioengineering*, 2(1):157–180.

[Vidaurre and Blankertz, 2010] Vidaurre, C. and Blankertz, B. (2010). Towards a cure for BCI illiteracy. *Brain Topography*, 23(2):194–198.

[Vidaurre et al., 2008] Vidaurre, C., Schlögl, A., Blankertz, B., Kawanabe, M., and Müller, K.-R. (2008). Unsupervised adaptation of the lda classifier for brain-computer interfaces. In *Proceedings of the 4th International Brain-Computer Interface Workshop and Training Course*, volume 2008, pages 122–127.

[Vourkas et al., 2000] Vourkas, M., Micheloyannis, S., and Papadourakis, G. (2000). Use of ann and hjorth parameters in mental-task discrimination. In *Advances in Medical Signal and Information Processing, 2000. First International Conference on (IEE Conf. Publ. No. 476)*, pages 327–332.

[Šidlauskas and Martišius, 2013] Šidlauskas, K. and Martišius, I. (2013). EEG duomenų klasifikavimas naudojant balsavimo ir daugiasluoksnius perceptronus. In *Informacinės technologijos : 17-oji tarpuniversitetinė magistrantų ir doktorantų konferencija : konferencijos pranešimų medžiaga*, volume 20, pages 55–58. Vilniaus universiteto Kauno humanitarinis fakultetas, Kauno technologijos universitetas, Vytauto Didžiojo universitetas.

[Vuckovic, 2010] Vuckovic, A. (2010). Motor imagery questionnaire as a method to detect BCI illiteracy. In *Applied Sciences in Biomedical and Communication Technologies (ISABEL), 2010 3rd International Symposium on*, pages 1–5.

[Wang et al., 2009a] Wang, B., Wan, F., Mak, P. U., Mak, P.-I., and Vai, M. I. (2009a). EEG signals classification for brain computer interfaces based on gaussian process classifier. In *Information, Communications and Signal Processing, 2009. ICICS 2009. 7th International Conference on*, pages 1–5.

[Wang et al., 2009b] Wang, C., Phua, K. S., Ang, K. K., Guan, C., Zhang, H., Lin, R., Sui Geok Chua, K., Ang, B. T., and Kuah, C. (2009b). A feasibility study of non-invasive motor-imagery BCI-based robotic rehabilitation for stroke patients. In *Neural Engineering, 2009. NER '09. 4th International IEEE/EMBS Conference on*, pages 271–274.

[Wang et al., 2010] Wang, J., Xu, G., Wang, L., and Zhang, H. (2010). Feature extraction of brain-computer interface based on improved multivariate adaptive autoregressive models. In *Biomedical Engineering and Informatics (BMEI), 2010 3rd International Conference on*, volume 2, pages 895–898.

[Wang et al., 2011] Wang, L., Xu, G., Wang, J., Yang, S., Guo, L., and Yan, W. (2011). Ga-svm based feature selection and parameters optimization for BCI research. In *Natural Computation (ICNC), 2011 Seventh International Conference on*, volume 1, pages 580–583.

[Wang et al., 2012] Wang, L., Xu, G., Yang, S., Guo, M., Yan, W., and Wang, J. (2012). Motor imagery BCI research based on sample entropy and svm. In *Electromagnetic Field Problems and Applications (ICEF), 2012 Sixth International Conference on*, pages 1–4.

[Whitney, 1937] Whitney, H. (1937). Topological properties of differentiable manifolds. *Bulletin of the American Mathematical Society*, 43(12):785–805.

[Williams and Peng, 1990] Williams, R. J. and Peng, J. (1990). An efficient gradient-based algorithm for on-line training of recurrent network trajectories. *Neural Computation*, 2(4):490–501.

[Witte, 2005] Witte, E. F. I. H. (2005). *Data Mining. Practical Machine Learning Tools and Techniques.*

[Wolpaw et al., 2002] Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., and Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clinical neurophysiology*, 113(6):767–791.

[Wright et al., 1990] Wright, J., Kydd, R., and Sergejew, A. (1990). Autoregression models of EEG. *Biological Cybernetics*, 62(3):201–210.

[Wu et al., 2014] Wu, W., Chen, Z., Gao, X., Li, Y., Brown, E., and Gao, S. (2014). Probabilistic common spatial patterns for multichannel EEG analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, PP(99):1–1.

[Wu et al., 2005] Wu, W., Gao, X., and Gao, S. (2005). One-versus-the-rest(ovr) algorithm: An extension of common spatial patterns(csp) algorithm to multi-class case. In *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, pages 2387–2390.

[Xu and Zhai, 2011] Xu, H. and Zhai, G. (2011). Ecg data compression based on wave atom transform. In *Multimedia Signal Processing (MMSP), 2011 IEEE 13th International Workshop on*, pages 1–5.

[Xu et al., 2014] Xu, Y., Yu, J., and Zhang, Y. (2014). Knn-based weighted rough t-twin support vector machine. *Knowledge-Based Systems*, 71:303–313.

[Yaacob et al., 2013] Yaacob, H., Abdul, W., and Kamaruddin, N. (2013). Classification of EEG signals using mlp based on categorical and dimensional perceptions of emotions. In *Information and Communication Technology for the Muslim World (ICT4M), 2013 5th International Conference on*, pages 1–6.

[Yamanaka and Yamamoto, 2010] Yamanaka, K. and Yamamoto, Y. (2010). Single-trial EEG power and phase dynamics associated with voluntary response inhibition. *Cognitive Neuroscience, Journal of*, 22(4):714–727.

[Yanagisawa et al., 2012] Yanagisawa, K., Sawai, H., and Tsunashima, H. (2012). Development of nirs-BCI system using perceptron. In *Control, Automation and Systems (ICCAS), 2012 12th International Conference on*, pages 1531–1535.

[Ying and Yusen, 2009] Ying, Y. and Yusen, W. (2009). New threshold and shrinkage function for ecg signal denoising based on wavelet transform. In *Bioinformatics and Biomedical Engineering, 2009. ICBBE 2009. 3rd International Conference on*, pages 1–4. IEEE.

[Yitembe et al., 2011] Yitembe, B., Crevecoeur, G., Van Keer, R., and Dupre, L. (2011). EEG inverse problem solution using a selection procedure on a high number of electrodes with minimal influence of conductivity. *Magnetics, IEEE Transactions on*, 47(5):874–877.

[Yuan et al., 2008] Yuan, Y., Li, Y., and Mandic, D. P. (2008). A comparison analysis of embedding dimensions between normal and epileptic EEG time series. *The Journal of Physiological Sciences*, 58(4):239–247.

[Zabidi et al., 2012] Zabidi, A., Mansor, W., Lee, Y. K., and Fadzal, C. (2012). Short-time fourier transform analysis of EEG signal generated during imagined writing. In *System Engineering and Technology (ICSET), 2012 International Conference on*, pages 1–4.

[Zander et al., 2010] Zander, T. O., Gaertner, M., Kothe, C., and Vilimek, R. (2010). Combining eye gaze input with a brain–computer interface for touchless human–computer interaction. *Intl. Journal of Human–Computer Interaction*, 27(1):38–51.

[Zhang et al., 2007] Zhang, J.-C., Xu, Y.-Q., and Yao, L. (2007). P300 detection using boosting neural networks with application to BCI. In *Complex Medical Engineering, 2007. CME 2007. IEEE/ICME International Conference on*, pages 1526–1530.

[Zschokke, 1995] Zschokke, S. (1995). *Klinische Elektroenzephalographie*. Springer-Verlag.

PUBLICATION LIST BY THE AUTHOR

Indexed in the Web of Science with Impact Factor

1. Martišius, Ignas; Damaševičius, Robertas; Jusas, Vacius; Birvinskas, Darius. Using higher order nonlinear operators for SVM classification of EEG data // Elektronika ir elektrotechnika = Electronics and Electrical Engineering. Kaunas: KTU. ISSN 1392-1215. 2012, nr. 3(119), p. 99-102. [Science Citation Index Expanded (Web of Science); INSPEC; Computers & Applied Sciences Complete; Central & Eastern European Academic Source]. [0,250]. [IF (E): 0,411 (2012)]
2. Birvinskas, Darius; Jusas, Vacius; Martišius, Ignas; Damaševičius, Robertas. Data compression of EEG signals for artificial neural network classification // Informacinės technologijos ir valdymas = Information technology and control / Kauno technologijos universitetas. Kaunas: KTU. ISSN 1392-124X. 2013, T. 42, nr. 3, p. 238-241. [Science Citation Index Expanded (Web of Science); INSPEC]. [0,354]. [IF (E): 0,813 (2013)]
3. Damaševičius, Robertas; Martišius, Ignas; Jusas, Vacius; Birvinskas, Darius. Fractional delay time embedding of EEG signals into high dimensional phase space // Elektronika ir elektrotechnika = Electronics and electrical engineering. Kaunas: KTU. ISSN 1392-1215. 2014, Vol. 20, no. 8, p. 55-58. [Science Citation Index Expanded (Web of Science); Inspec; Computers & Applied Sciences Complete; Central & Eastern European Academic Source; Scopus]. [0,250]. [IF (E): 0,561 (2014)]
4. Birvinskas, Darius; Jusas, Vacius; Martišius, Ignas; Damaševičius, Robertas. Fast DCT algorithms for EEG data compression in embedded systems // Computer science and information systems. Novi Sad: University of Novi Sad. ISSN 1820-0214. 2015, Vol. 12, Iss. 1, p. 49-62. [Science Citation Index Expanded (Web of Science)]. [0,250]. [IF (E): 0,477 (2014)]

Indexed in the Web of Science without Impact Factor

1. Martišius, Ignas; Vasiljevas, Mindaugas; Šidlauskas, Kęstutis; Turčinas, Rūtenis; Plauska, Ignas; Damaševičius, Robertas. Design of a neural interface based system for control of robotic devices // Information and software technologies : 18th international conference, ICIST 2012, Kaunas, Lithuania, September 13-14, 2012 : proceedings / [Edited by] Tomas Skersys, Rimantas Butleris, Rita Butkiene. Berlin: Springer, 2012. (Communications in computer and information science, Vol. 319, ISSN 1865-0929), ISBN 9783642333071. p. 297-311. [Conference Proceedings Citation Index]. [0,167]
2. Martišius, Ignas; Damaševičius, Robertas. Class-adaptive denoising for EEG data classification // Artificial intelligence and soft computing : proceedings of the 11th international conference, ICAISC 2012, April 29 - May 3, 2012, Zakopane, Poland, Part 2. Berlin: Springer-Verlag, 2012. (Lecture Notes in Computer Science, Vol. 7268, ISSN 0302-9743), ISBN 978642293498. p. 302-309. [Conference Proceedings Citation Index; SpringerLINK]. [0,500]
3. Birvinskas, Darius; Jusas, Vacius; Martišius, Ignas; Damaševičius, Robertas. EEG dataset reduction and feature extraction using discrete cosine transform // UKSim-AMSS EMS 2012 : 6th European Modelling Symposium on Mathematical Modeling and Computer Simulation 2012, 14-16 November, 2012, Malta : proceedings. Los Alamitos, Washington, Tokyo: IEEE Computer Society, 2012, ISBN 9780769549262. p. 199-204. [IEEE]. [0,250]
4. Martišius, Ignas; Birvinskas, Darius; Damaševičius, Robertas; Jusas, Vacius. EEG dataset reduction and classification using wave atom transform // Artificial neural networks and machine learning - ICANN 2013 : 23rd international conference on artificial neural networks, Sofia, Bulgaria, September 10-13, 2013 : proceedings / [edited by] Valeri Mladenov, Petia Koprinkova-

Hristova, Guenther Palm. New York: Springer, 2013. (Lecture notes in computer science, Vol. 8131, ISSN 0302-9743), ISBN 9783642407277. p. 208-215. [Conference Proceedings Citation Index; SpringerLINK]. [0,250]

5. Martišius, Ignas; Šidlauskas, Kęstutis; Damaševičius, Robertas; Damaševičius, Robertas. Real-time training of voted perceptron for classification of EEG data // International Journal of Artificial Intelligence. Kaunas: KTU. ISSN 0974-0635. 2013, Vol. 10, iss. S13, p. 41-50. [INSPEC; SCOPUS; Zentralblatt MATH; IndexCopernicus]. [0,250]

Articles in conference proceedings

1. Birvinskas, Darius; Martišius, Ignas. Aparatinė dvimatės DCT ir IDCT realizacija, naudojant Bindct algoritmą // Informacinės technologijos : 16-oji tarpuniversitetinė magistrantų ir doktorantų konferencija : konferencijos pranešimų medžiaga / Kauno technologijos universitetas, Vytauto Didžiojo universitetas, Vilniaus universiteto Kauno humanitarinis fakultetas. Kaunas: Technologija. ISSN 2029-249X. 2011, p. 129-132. [0,500]
2. Šidlauskas, Kęstutis; Martišius, Ignas. EEG duomenų klasifikavimas naudojant balsavimo ir daugiasluoksnius perceptronus // Informacinės technologijos : 17-oji tarpuniversitetinė magistrantų ir doktorantų konferencija : konferencijos pranešimų medžiaga / Vilniaus universiteto Kauno humanitarinis fakultetas, Kauno technologijos universitetas, Vytauto Didžiojo universitetas. Vilnius: Vilniaus universitetas. ISSN 2029-249X. 2012, p. 39-42. [0,500]
3. Vasiljevas, Mindaugas; Turčinas, Rūtenis; Martišius, Ignas. Netiesinių operatorių taikymas EEG duomenų apdorojimui // Informacinės technologijos : 17-oji tarpuniversitetinė magistrantų ir doktorantų konferencija : konferencijos pranešimų medžiaga / Vilniaus universiteto Kauno humanitarinis fakultetas, Kauno technologijos universitetas, Vytauto Didžiojo universitetas. Vilnius: Vilniaus universitetas. ISSN 2029-249X. 2012, p. 23-26. [0,333]
4. Martišius, Ignas. EEG signal processing methods for BCI applications // XV International PhD Workshop OWD 2013 : under the auspices of deans of electrical, electronic and computer science faculties of engineering, IEE - The Institution of Engineering and Technology, IEEE - Institute of Electrical and Electronics Engineers - Polish Section, 19-22 October 2013, Wisła, Poland. Gliwice: Organizing Committee of the Symposium PPEE & Seminar BSE, 2013. (Archiwum konferencji PTETiS = Conference arcives PTETiS, Vol. 33), ISBN 9788393542727. p. 84-89. [1,000]
5. Vasiljevas, Mindaugas; Martišius, Ignas; Šumskas, Tomas. Evaluation of user fatigue in neural computer interface system // Informacinės technologijos : 19-oji tarpuniversitetinės magistrantų ir doktorantų konferencija "Informacinė visuomenė ir universitetinės studijos" (IVUS 2014) : konferencijos pranešimų medžiaga / Kauno technologijos universitetas, Vytauto Didžiojo universitetas, Vilniaus universiteto Kauno humanitarinis fakultetas. Kaunas: Technologija. ISSN 2029-249X. 2014, p. 193-196. [0,333]
6. Vasiljevas, Mindaugas; Martišius, Ignas; Šumskas, Tomas. Evaluation of user fatigue in neural computer interface system // Informacinės technologijos : 19-oji tarpuniversitetinė magistrantų ir doktorantų konferencija "Informacinė visuomenė ir universitetinės studijos" (IVUS 2014) : konferencijos pranešimų medžiaga / Kauno technologijos universitetas, Vytauto Didžiojo universitetas, Vilniaus universiteto Kauno humanitarinis fakultetas. Kaunas: Technologija. ISSN 2029-4832. 2014, p. 193-196. [0,333]

SL344. 2016-03-03, 19 leidyb. apsk. l. Tiražas 12 egz. Užsakymas 101.

Išleido Kauno technologijos universitetas, K. Donelaičio g. 73, 44249 Kaunas
Spausdino leidyklos „Technologija“ spaustuvė, Studentų g. 54, 51424 Kaunas