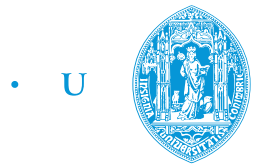


University of Coimbra
Faculty of Science and Technology
Department of Electrical and Computer Engineering

Biosignal Classification for Human Interface with Devices and Surrounding Environment

Gabriel Pereira Pires

2011

Biosignal Classification for Human Interface with Devices and Surrounding Environment

Gabriel Pereira Pires

Submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

Department of Electrical and Computer Engineering
University of Coimbra, Portugal

under supervision of

Prof. Dr. Urbano Nunes (advisor)
Prof. Dr. Miguel Castelo-Branco (co-advisor)

To my close family

Acknowledgments

This work is the outgrowth of research carried out at the Institute for Systems and Robotics (ISR) of the University of Coimbra in the last five years. When I started in late 2006, the research on 'brain-computer interfaces' was a new experience to me, and an unexplored research area in ISR. To get here today, I received many direct and indirect contributions that helped me to overcome the challenges of this magnificent research area.

Undoubtedly, my first thanks go to my supervisors, Prof. Urbano Nunes and Prof. Miguel Castelo-Branco. Prof. Urbano Nunes supervised me during my B.Sc, M.Sc and Ph.D. I owe him a lot of my training as a student and researcher. During this Ph.D work, Prof. Urbano's supervision was outstanding, by his methodological contributions, fruitful discussions, continuous motivation, friendship, and for offer me all the necessary means to carry out this work. He gave me also the freedom to follow my own directions. To Prof. Miguel Castelo-Branco, I want to thank him for the fruitful discussions, for his continued support, friendship, enthusiastic and outstanding way to motivate me, and for introducing me important neuroscience concepts.

During my Ph.D work, 48 people participated in the experiments, from whom 14 suffer from motor impairments. I'm deeply thanked to all of them. Special thanks go to Charlotte, my wife, and Paulo Sergio, my best friend, who participated in almost all pilot experimental sessions of the proposed paradigms, for many hours. Their time, patience and contributions were crucial to eliminate implementation bugs, and to rightly design the visual paradigms.

From the Associação de Paralisia Cerebral de Coimbra (APCC), where part of the experiments took place, I want to thank to Dr. Carlos Barata, from the board direction, who gave me the entire support for experimental tests. I would like to thank the indispensable assistance of occupational therapists Andreia, Diana, Paula and Vera, during the experimental sessions, and for their help to select the APCC participants.

From the Hospitals of the University of Coimbra (HUC), where part of the experiments took place, I want to thank to Dr. Luis Negrão, who gave me the entire support for experimental tests, and who selected ALS patients and rated their functional level.

I owe a lot to ISR to have provided excellent conditions and resources that allowed me to accomplish my Ph.D work. I want to thank also to the Biomedical Institute for Research in Light and Image (IBILI) for providing me all the conditions and resources for experimental tests. I am very grateful to Instituto Politécnico de Tomar (IPT), where I teach as Prof. Adjunto, for giving me the conditions to conciliate the lecturing activities with this Ph.D work, during the past five years.

I acknowledge two partial fellowships, one from Fundação para a Ciência e a Tecnologia (FCT) under Grant SFRH/BD/29610/2006, and the other from a PROTEC research fellowship under Grant SFRH/BD/49881/2009. This work has been in part supported by Fundação para a Ciência e Tecnologia (FCT), under Grants PTDC/EEA-ACR/72226/2006 (PMITS06) and RIPD/ADA/109661/2009 (INTERFACE10).
I am also very grateful to my colleagues from ISR, Cristiano Premebida, Luciano Oliveira and Oswaldo Ludwig, with who I had some fruitful discussions about machine learning. To Ana Lopes, Ana Vieira and Pedro Correia, my friends and colleagues at IPT, my thanks for their support. To all my friends, in particular, Paulo Sérgio, José Beato and Marco Monteiro, thanks for their concern and motivation.

Finally, I would like to express my gratitude to my wife Charlotte, my daughters Lara and Valentina, my mother Piedade, my father Gabriel, and my sisters Alcina and Isonda, for their concern and constant encouragement and support.

Abstract

Brain computer interface (BCI) is emerging as a new communication channel for individuals with severe motor disabilities. Yet, its effective application in real-world scenarios is still limited by low communication rates and by other practical aspects. Notwithstanding, the constant improvement on techniques for brain activity recording provides signals with increasing signal-to-noise ratios, and the increase in computational power allows the use of more complex signal processing techniques, in real-time. In electroencephalographic (EEG) based BCIs, the recording technique used in this work, much can be done to improve the performance of BCI systems and to overcome practical limitations, moving toward its effective use in clinical and non-clinical applications. Therefore, the aim of this thesis is to contribute with new signal processing methods and new visual paradigms in the context of BCIs based on a neurophysiologic signal designated by 'P300 event related potential'. Moreover, several related practical issues are addressed.

In particular, the improvement of communication rates has been accomplished by two main approaches: 1) development of signal processing techniques based on statistical spatial filtering; and 2) design of new visual paradigms/protocols. Improving these two components simultaneously, significantly increases the accuracy and transfer rates of P300 BCIs. Regarding the development of statistical spatial filtering, several methods were proposed in the time domain and in the frequency domain. The frequency domain approach led to a spatio-spectral method combining simultaneously spatial and temporal filtering. The performance of the proposed methods was assessed offline and online though extensive experiments performed by able-bodied individuals and individuals with severe motor disabilities. The results demonstrated the effectiveness of the proposed methods. Regarding the design of new paradigms, a lateral single character (LSC) speller was proposed to overcome several of the limitations of the standard row-column speller. Additionally, it introduced a new neurophysiologic dimension besides P300, which is related to the left vs. right visual presentation layout in the design of the paradigm. The effective improvements regarding communication rates and P300 enhancement were demonstrated by online experiments performed by able-bodied individuals and individuals with severe motor disabilities.

Several other practical aspects were addressed, through the development of a gaze independent P300-based BCI, and through the development of an asynchronous, subject controlled, P300-based BCI. The gaze independent BCI makes possible its control without moving the eyes, an important issue for individuals with severe ocular limitations. The asynchronous BCI was proposed in the context of a robotic wheelchair steering, allowing

an automatic subject dependent switch On/Off of the BCI.

We hope that this work has pointed to new directions on the application of statistical spatial filters. Concerning visual paradigms, we hope that the novel proposed approaches will be adopted and improved by more research groups.

Resumo

A interface cérebro-computador (ICC) está a emergir como um novo canal de comunicação para indivíduos com deficiências motoras graves. Ainda assim, a sua efectiva aplicação em cenários do mundo real é limitada pelas baixas taxas de comunicação e por outros aspectos práticos. Não obstante, a constante melhoria das técnicas de registo da actividade cerebral fornece sinais com crescentes relações sinal-ruído, e o aumento da capacidade computacional permite o uso de técnicas de processamento de sinal mais complexas em tempo real. Em ICCs baseadas em eletroencefalografia (EEG), a técnica de registo utilizada neste trabalho, existem muitas formas de melhorar o desempenho de sistemas ICC, e de superar limitações práticas, caminhando assim em direcção à sua efectiva utilização em aplicações clínicas e não-clínicas. Sendo assim, esta tese pretende contribuir com novos métodos de processamento de sinal e com novos paradigmas visuais, no contexto de ICCs baseadas num sinal neurofisiológico designado 'potencial evocado P300'. Para além disso, são abordadas várias questões práticas relacionadas.

Em particular, conseguiu-se melhorar as taxas de comunicação através de duas abordagens principais: 1) desenvolvimento de técnicas de processamento de sinal baseadas em filtragem espacial estatística; e 2) concepção de novos paradigmas visuais / protocolos. A melhoria simultânea deste dois aspectos permite aumentar significativamente a precisão e as taxas de transferência de ICCs baseadas em P300. Quanto ao desenvolvimento de filtragem espacial estatística, foram propostos vários métodos no domínio do tempo e no domínio da frequência. A abordagem no domínio da frequência levou a um método espaço-espectral que combina filtragem espacial e temporal. O desempenho dos métodos propostos foi avaliado *offline* e *online* num extenso conjunto de experiências realizadas por indivíduos saudáveis e indivíduos portadores de deficiência motora grave. Os resultados demonstraram a eficácia dos métodos propostos. Em relação à concepção de novos paradigmas, foi proposto um soletrador designado por "lateral single character" (LSC), para superar algumas das limitações do soletrador padrão de "linhas e colunas". Além disso, este paradigma introduziu uma nova dimensão neurofisiológica para além do P300, a qual está relacionada com a concepção esquerda vs. direita do paradigma em termos do *layout* do estímulo visual. A efectiva melhoria das taxas de comunicação e do sinal P300 foi demonstrada por experiências *online* realizadas por indivíduos saudáveis e indivíduos portadores de deficiência motora grave.

Foram abordados vários outros aspectos práticos, através do desenvolvimento de uma ICC baseada em P300, independente do olhar do utilizador, e através do desenvolvimento de uma ICC baseada em P300 com funcionamento assíncrono, dependente da intenção

do participante. A ICC independente do olhar torna possível o seu controlo sem mover os olhos, uma questão importante para indivíduos com limitações oculares graves. A ICC assíncrona foi proposta no contexto da condução de uma cadeira de rodas robótica, permitindo um modo automático de ligar/desligar a ICC.

Esperamos que este trabalho tenha apontado novas direcções na aplicação de filtros espaciais estatísticos. Relativamente aos paradigmas visuais, esperamos que as novas abordagens propostas sejam adoptadas e aperfeiçoadas por mais grupos de investigação.

Contents

List of Tables

List of Figures

List of Abbreviations

ADHD	Attention deficit hyperactivity disorder
ALS	Amyotrophic lateral sclerosis
ALSFRS	ALS functional rating scale
AR	Auto-regressive
ASSFCB	Automatic spatio-spectral Fisher criterion beamformer
BCI	Brain-computer interface
BMI	Brain-machine interface
BOLD	Blood oxygen level dependent
C-FMS	Cascade Fisher and maximum signal-to-noise ratio
CAR	Common average reference
CP	Cerebral palsy
CSP	Common spatial patterns
CSTP	Common spatio-temporal patterns
DFT	Discrete Fourier transform
DMD	Duchenne muscular dystrophy
ECoG	Electrocorticography
EEG	Electroencephalography
EMG	Electromyography
EOG	Electrooculography
ERD	Event related desynchronization
ERP	Event related potential
ERS	Event related synchronization

FC	Fisher criterion
FES	Functional electric stimulation
FFT	Fast Fourier transform
FLD	Fisher's linear discriminant
FMRI	Functional magnetic resonance imaging
FNR	False negative rate
FPR	False positive rate
GEVD	Generalized eigenvalue decomposition
GIBS	Gaze independent block speller
ICA	Independent component analysis
IDFT	Inverse discrete Fourier transform
ISI	Inter-stimulus interval
ITI	Inter-trial interval
ITR	Information transfer rate
LAP	Laplacian
LDA	Linear discriminant analysis
LFP	Local field potentials
LIS	Locked-in state
LSC	Lateral single character
MEG	Magnetoencephalography
MND	Motor neuron disease
mRMR	Minimum redundancy maximum relevance
MRP	Movement related potential

NB	Naïve Bayes
NIRS	Near-infrared spectroscopy
NN	Neural networks
PBR	Practical bit rate
PCA	Principal component analysis
qNB	Quadratic naïve Bayes
RC	Row-column
RP	Readiness potential
SCI	Spinal cord injury
SCP	Slow cortical potential
SFS	Sequential feature selection
SMR	Sensorimotor rhythms
SNR	Signal-to-noise ratio
SOA	Stimulus onset asynchrony
SPM	Symbols per minute
SSFCB	Spatio-spectral Fisher criterion beamformer
SSVEP	Steady-state visual evoked potential
SVM	Support vector machines
SWLDA	Stepwise linear discriminant analysis
TNR	True negative rate
TPR	True positive rate
VEP	Visual evoked potential

Chapter 1

Introduction

Contents

"Knowledge is our destiny."
Carl Sagan, in "The Dragons of Eden"

1.1 Motivation

Human communication and interaction with surrounding environment are two of the most important things for the human being on which depends its autonomy and survival. Individuals who suffer from severe motor disabilities by neuromotor disorders, such as amyotrophic lateral sclerosis (ALS) and cerebral palsy (CP), have their major communication abilities affected. Some of them suffer from total lack of motor control or very low dexterity affecting head, limbs, eyes and speech. In neuromotor diseases such as ALS, also known as Lou Gehrig's disease, there is a progressive degeneration of lower and upper motor neurons that lead to a progressive muscular weakness. Affected individuals may ultimately lose the ability to control any voluntary muscle entering in a "locked-in" state.

Brain computer interfaces (BCIs) emerge as a feasible type of human-computer and human-machine interfaces that open a new communication channel to those persons, without recurring to the conventional output pathways of peripheral nerves and muscles [Wolpaw 2002]. The goal of a BCI is to translate brain activity, that has been volitionally

induced by the user, into commands that can be used to control a computer or other input devices. Individuals in a "locked-in" condition are conscious and aware of the surrounding environment and thus may use, through a BCI, brain activity to communicate. For individuals who still retain a weak ability to control some part of the body, but whose motor performance is very low, showing difficulty on controlling some standard interfaces, such as as speech recognition, eye tracking, head tracking, and head or teeth switches, BCI can be used as an alternative interface. Increasing the communication abilities of these individuals by improving the accessibility to computers and means of locomotion such as wheelchairs, gives them a higher level of independence and self-satisfaction, thus enhancing their quality of life. Besides "assistive communication", BCI finds many other applications seeking restoration or repair of human cognitive or sensory-motor functions. Examples include the clinical applications such as the treatment of neurobehavioral disorders such as attention-deficit and hyperactivity disorder (ADHD) in children and adolescents, treatment of seizure disorders, neuro-rehabilitation for recovery from stroke and traumatic brain injuries, and entertainment in game applications for healthy users.

From the first pioneering works of Vidal [Vidal 1973], who first introduced the term "Brain-Computer Interface", many BCI systems have been developed in particular in the last decade. Nevertheless, today, the use of BCI is still very limited to experimental tests, although some patients already use BCI regularly in their daily lives as a communication device. Several issues have to be addressed to move toward a practical and effective BCI system. Some of theses issues are addressed in the next section.

1.2 Challenges of EEG-based BCIs

BCI is a complex research area that integrates knowledge from different fields of neuroscience, biophysics, psychology, computer science, and electrical engineering. Effective BCI promises depend on the contribution and combination of all these interdisciplinary fields, which all still face several major challenges. The BCI systems implemented in this thesis are based on a non-invasive method that measures electroencephalographic (EEG) activity from the scalp. One of the main challenges of EEG-based BCIs is to increase their current low communication transfer rates (a few symbols per minute). EEG is characterized by a very low signal-to-noise ratio (SNR). Thus, the decoding of user intentions from brain patterns requires the application of signal processing and pattern recognition techniques that can enhance the desired components and attenuate noise from EEG data. The low SNR is due to several reasons, namely, the low amplitude of the signals, the

poor spatial resolution and the high susceptibility to muscular and electrical artifacts. Moreover, EEG signals are nonstationary and present inter-subject and within-subject variability. Thus, the BCI must be calibrated individually whenever it is used. Several usability issues are also relevant: BCI requires more attention and concentration than standard interfaces; and BCI based on external stimuli may cause fatigue and discomfort. These two issues limit longtime use of BCI. Moreover, some types of BCIs may require long-term learning. Other aspects such as setup time needed to place the electrodes, due to gel preparation and significant number of channels, and the unpractical and unnatural electrode caps currently available are also difficulties to BCI use. Although some dry electrodes are already available, providing a fast setup, these electrodes still exhibit lower SNR when compared with wet electrodes (with gel). Finally, most of current BCIs still work in synchronous mode, that is, the time periods for which the user wishes to control the BCI system is determined by the application and not volitionally by the user. In certain real-world scenarios (e.g., control of robotic devices), this limitation derails its use, or restricts the use of BCI to permanent supervision of humans.

1.3 Goals and key contributions

There are currently three major neural mechanisms used in EEG-based BCIs: control of sensorimotor rhythms, visual evoked potentials and P300 event-related potentials. This thesis focuses on the implementation of visual P300-based BCIs toward their effective application in real-world scenarios. The implementation of a P300-based BCI can generically be divided in two parts: development of algorithms for signal processing and classification; and design of the paradigm/protocol. By improving simultaneously these two parts it is possible to greatly increase the accuracy and transfer rates of P300 BCIs. Moreover, the simultaneous combination of different neurophysiologic features, induced by novel paradigm strategies is a major contribute to increase performance accuracies. Other relevant aspects for the practical application of P300-based BCIs were also addressed, namely the reduction of calibration time, and the asynchronous control of BCI. The methodologies proposed were assessed online and offline in several case-studies with participation of able-bodied individuals and individuals with motor disabilities suffering from disorders such as amyotrophic lateral sclerosis (ALS), Duchenne muscular dystrophy (DMD), cerebral palsy (CP) and spinal cord injury (SCI).

The following specific contributions were given in this thesis:

- **Statistical spatial filters.** The improvement of signal processing and classification

methodologies are key issues to increase the transfer rate of P300-based BCIs. While most of the research in this topic has been on the development of new classifiers, we researched the application of optimum spatial filters (in the statistical sense). With this approach, spatially filtered signals show a remarkable increase in SNR, which improves classification performance. Taking as starting point a Max-SNR beamformer, two spatial filters were implemented, a beamformer based on the Fisher criterion (FC) and a spatial filter, the C-FMS method, combining sub-optimally FC and Max-SNR beamformers [Pires 2011d]. In experiments with able-bodied and disabled individuals controlling a Row-Column (RC) BCI speller, the C-FMS method led to performance results above state-of-the-art. Additionally, FC beamformer was extended from the time domain to the frequency domain, leading to a spatio-spectral Fisher criterion beamformer (SSFCB). The combination of spatial filtering with temporal filtering increased flexibility and classification performance.

- **Novel visual paradigms.** Visual paradigm design greatly influences the quality of the evoked P300 potential. A new proposed paradigm, called lateral single-character (LSC) speller, enhances the P300 component and increases classification accuracy. The paradigm was compared with the standard RC speller, through extensive online experimental tests performed by able-bodied and disabled individuals [Pires 2011b]. LSC showed an increased transfer rate when compared to RC with results above state-of-the-art. Moreover, the paradigm design of LSC allows the combination of P300 and SSVEP features, introducing a new neurophysiologic dimension related to left/right stimuli presentation. Thus, LSC can still have a margin for improvement.

- **Gaze-independent visual paradigm.** In the last few years, the importance of gazing a target stimulus to effectively control a P300-based BCI has been discussed. This gaze dependence/independence issue is very relevant when the interface intends to be controlled by individuals with severe limitations on ocular movements (e.g., advanced stages of ALS). This issue was addressed by proposing a two-level gaze independent paradigm called GIBS (gaze independent block speller) [Pires 2011c]. Experimental tests showed that GIBS can be used without moving the eyes and with effective transfer rates. A generic bit rate metric was proposed for two-level paradigms, allowing bandwidth comparison with other paradigms.

- **Asynchronous BCI control.** A P300-based paradigm designed specifically to steer a wheelchair was initially implemented [Pires 2008b]. We realized that the effective application of this interface required an asynchronous control of the BCI, i.e.,

user had to be able to automatically switch On/Off the interface. With this purpose, a 3-state paradigm was researched and implemented based on several approaches.

- **P300-BCI assessment with cerebral palsy individuals.** For the best of our knowledge, this thesis is the only work to date where P300-based BCIs were tested by a significant group of individuals suffering from cerebral palsy. This allowed us to reach some preliminary results and assessment about the effective application of BCI in this particular group.

As already referred, the requirement of a calibration of the BCI before each experimental session is a limitative factor. We limited the calibration time of all proposed P300-based BCIs to about 5 minutes, far below most current systems. Moreover, the maximum number of electrodes was limited to 12, providing a realistic scenario.

Two additional contributions have also been given. The first was the development of a Tetris game controlled by BCI, that can be adapted as a neurotherapy tool for children with neurobehavioural disorders [Pires 2011a]. The second is an automatic detection of P300 components evoked by subliminar visual tasks, which can be used in clinical applications for vision [Teixeira 2011].

1.4 Thesis outline

The remainder of this thesis is organized as follows.

- **Chapter 2.** This chapter presents neurophysiologic background for basic understanding of our work. Moreover, a short state-of-the-art allows the reader to frame the contributions of this thesis.

- **Chapter 3.** This chapter provides all technical and scientific aspects related to the implementation and assessment of a P300-based BCI. It gives the necessary background to understand the proposed methodologies in remaining chapters.

- **Chapter 4.** Basic description of classification methods is presented here. The chapter describes methods effectively implemented offline or online along the thesis.

- **Chapter 5.** This is a core chapter. It formally presents the proposed spatial filters. The methods are extensively assessed offline and online through a case-study with able bodied and motor disabled participants. The methods are also assessed with benchmark datasets.

- **Chapter 6.** A new paradigm is presented: the LSC speller. Aspects designed to eliminate typical limitations of P300-based spellers are described. The paradigm is compared with the RC speller through extensive experimental tests performed by able bodied and motor disabled participants.

- **Chapter 7.** This chapter presents the design and assessment of GIBS, a speller addressing the gaze independence issue in P300-based BCI. Additionally, a new metric is obtained specifically for this type of paradigm.

- **Chapter 8.** A paradigm specifically designed to steer a wheelchair is described. Then, some approaches for asynchronous control are described.

- **Chapter 9.** This chapter draws some conclusions and suggest some future directions.

- **Appendices C and D.** These appendices present some satellite case studies that were performed along this thesis with potential relevance for clinical application.

Chapter 2

Neurophysiologic background and state of the art

Contents

"I can't put it into words. It's just- I use my brain. I just thought it. I said, 'Cursor go up to the top right.' And it did, and now I can control it all over the screen. It will give me a sense of independence."

Matthew Nagle (C3 Tetraplegic, first person to use intracortical BCI in 2004)

This chapter introduces a BCI system, first starting by its definition and then generically describing its main functional elements. Recording methods, main neurophysiological signals used in BCI and current potential applications are presented providing the current state-of-the-art.

2.1 Introduction to BCI

2.1.1 Definition

The definition of brain-computer-interface (BCI) or brain-machine-interfaces (BMI) is not unique and has changed along the last decades. Currently, the definition most accepted

by the BCI community is the one given by Wolpaw [Wolpaw 2002]: *"A BCI is a communication system in which messages or commands that an individual sends to the external world do not pass through the brain's normal output pathways of peripheral nerves and muscles"*. For other words, user's intent information is encoded only by means of brain signals. The term 'neuroprosthesis' is sometimes used interchangeably with the same meaning of BCI, however, neuroprosthesis is a more generic term that differs from BCI in two aspects. It can interact with any part of the the nervous system, i.e., with central nervous system but also with peripheral nervous system (e.g. median nerve to control a prosthetic hand), and it can provide inputs to the nervous system (e.g. retinal implant or cochlear implant) [BrainAble 2010]. A BCI translates brain signals volitionally produced by the user. These signals can be used to control a computer or a robotic device (e.g. a wheelchair or a prosthetic device). Thus, BCI opens a new communication channel to those who suffer from severe motor disabilities, unable to control muscular activity. A generic view of a BCI system is presented in Fig. 2.1. According to the above BCI definition, BCIs can fall into two classes, dependent and independent [Wolpaw 2002]:

Dependent BCI A dependent BCI considers that besides brain activity, muscular activity is required to control the BCI. For example, in a BCI based on visual stimuli placed at different screen locations, if the user has to move the eyes to one of the stimulus to evoke a potential, then it is considered a dependent BCI. Despite the brain's output channel is EEG, the user requires extraocular muscle activity to gaze the stimulus. In this case, more efficient interfaces could be used to monitor eye movements (e.g., eye-tracking camera).

Independent BCI On the other hand, an independent BCI, sometimes referred as a 'true' BCI, does not depend in any way on peripheral nerves and muscles to generate the brain activity. For example, the BCI based on a visual P300 evoked potential is considered independent because the EEG signal depends only on the users's intent, not requiring gaze shifting.

People with severe motor impairments, lacking of all normal output pathways, may be unable to control a dependent BCI. The importance of gaze shifting depends obviously also on the design of the task and visual display, as well as on other exteroception and proprioception characteristics of users [Allison 2008]. The dependence/independence BCI classification is sometimes controversial or non-consensual, as it will be discussed in section 2.2. This issue will be addressed in detail in chapter 7.

2.1.2 Signal acquisition

Several methods exist for brain signal acquisition, ranging from non-invasive to highly invasive techniques. The biophysical nature of the signal source can be grouped in electrophysiological, magnetic and metabolic. A summary of these measurement techniques are briefly reviewed below comparing some of the main characteristics. Extended surveys can be found in [BMBF 2005, Mak 2009].

2.1.2.1 Electrophysiological signals

Electrophysiological signals are usually classified according to their degree of recording invasiveness. Three types are currently used in BCI: electroencephalographic (EEG), electrocorticographic (ECoG) and intracortical.

Electroencephalography EEG is one of the most popular non-invasive methods. It was measured for the fist time by Hans Berger in 1929 [Jung 1979]. These signals are recorded from the surface of the scalp. This method is very popular because the measurement devices are simple, low cost, safe, and the measurement preparation is relatively fast. Only an abrasive gel for skin cleaning and a conductivity gel are used to improve the conductivity between the scalp and the electrodes. These are usually placed at standard positions according to the 10-20 international system or extended positions of this standard (Fig. 3.6). However, EEG presents some limitations, namely, limited frequency range, low spatial resolution, and susceptibility to several types of artifacts. An EEG electrode measures the potentials resulting from the summation of the synapses of millions (ensemble) of neurons under the area beneath the surface of the electrode. However, due to volume conduction in the skull and scalp, the electrode receives also EEG signals from surrounding regions up to 10 cm [Srinivasan 1999]. This poor spatial resolution results on a mixing of brain components coming from different neurophysiologic sources. Furthermore, the skull attenuates signals (amplitudes in the order of $10 - 20\mu V$), in particular, in frequencies above 50 Hz. The presence of neurophysiologic artifacts, such as cranial muscular activity, eye movements and blinks, as well as external electromagnetic interferences, further degrade the quality of the signals. The combined effect of these noise sources leads to a recorded signal with a relative low signal-to-noise ratio (SNR).

Electrocorticographic ECoG is measured under the skull either outside the dura mater (epidural) or under the dura mater (subdural). Thus, the placement of the electrodes requires an intracranial invasive surgical procedure. The elimination of the insulating effect of the skull and scalp significantly improves the spatial resolution (tenths of millimeters), the attenuation effects (amplitudes in the order of $50 - 100\mu V$), and reduces

the presence of artifacts. This leads to a higher SNR and a wider frequency range (0-500 Hz) when compared to EEG recordings. Experiments with ECoG-BCIs have been made in patients with epilepsy who undergo the temporary placement of electrodes for monitoring and localization of seizure foci prior to surgical procedure [Schalk 2008]. The BCI experiments are made during the periods of monitoring, typically of 2-3 weeks. Studies in [Schalk 2004, Leuthardt 2006] show that patients are able to perform 2D control of a cursor in a computer screen with little amounting of training.

Intracorticography Intracortical recordings require the implantation of microelectrodes in the brain. They measure local field potentials (LFP) and neural action potentials (spikes). However, this highly invasive procedure carries serious risks of infections for patients. Furthermore, it is difficult to maintain stable long-term recordings and the technical requirements of single-neuron recordings are high. For these reasons, intracortical BCI have been studied mainly in monkeys [Carmena 2003] and very few in humans [Kim 2008]. Some experiments show monkeys performing reach-and-grasp tasks in 3D (three-dimension) by means of a 6DOF (degrees of freedom) robot arm [Carmena 2003]. In [Kim 2008], a tetraplegic human, using neural activity recorded from the motor cortex, acquired the ability to control a computer cursor in 2D (position and velocity).

Despite the great quality of the signals of invasive approaches, their current use in BCI for clinical applications is quite limited. Therefore, scalp recorded EEG remains one of the most used methods, presenting successful results in clinical applications and other areas as shown in the following sections and from our own achievements in this thesis.

2.1.2.2 Magnetic and Metabolic signals

Magnetoencephalography Magnetoencephalography (MEG) is a non-invasive method that detects magnetic fields induced by electrical currents in the active neurons. The magnetic field strength is very small ($< 10^{-12}$ Tesla) and therefore, extremely sensitive devices are necessary to detect them. MEG has a temporal resolution similar to EEG, but a higher spatial resolution and wider frequency range. It has been used in some experimental BCI studies [Mellinger 2007]. However, despite the attractive characteristics, MEG requires expensive bulky equipment and shielded environment, making impracticable its clinical use.

Functional Magnetic Resonance Imaging Functional magnetic resonance imaging (fMRI) measures metabolic activity through blood oxygen level dependent (BOLD) signal. Compared to EEG it has a better spatial resolution, however it presents a poor temporal resolution, in the order of seconds, i.e., BOLD signal is observed only a few seconds after

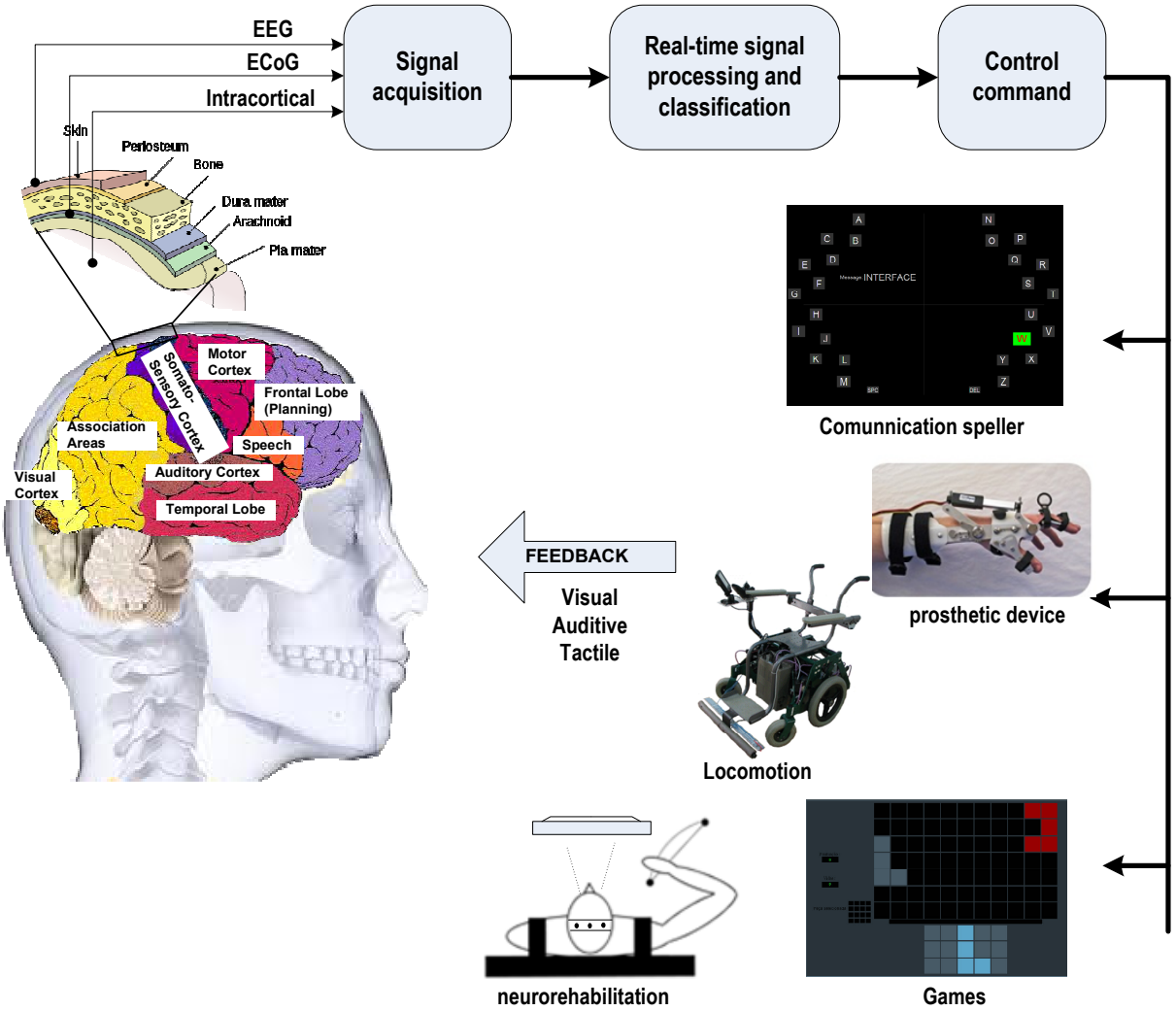


Figure 2.1: Generic representation of functional elements of a BCI system (figure inspired in [Mak 2009]). Signal recording from the surface of the scalp (EEG), intracranial over the cortical surface (ECoG), and within the brain (intracortical); Signals are processed and classified in real-time and then translated into control commands; These commands can be used for communication (e.g. speller devices), to control robotic devices for locomotion (e.g., wheelchair) or prosthetic devices (e.g., orthotic hand (Oregon Talon^{TM})), for game control (e.g., used as neurotherapy), and for neurorehabilitation (e.g., recovery of damaged brain regions). A BCI system always requires the feedback (e.g., visual, auditive or tactile) of the actions.

the neural activation. This temporal resolution strongly limits BCI applications. Notwithstanding, as a proof-of-concept, researchers successfully achieved real-time BCI operation [Weiskopf 2004].

Near-infrared spectroscopy Near-infrared spectroscopy (NIRS) is a recent non-invasive tech-

nique that measures the cortical activity through optical methods. Near-infrared light, light with a wave-length in the range of $700 - 1000$ nm, penetrates the scalp, suffering changes in light absorption and reflection according to BOLD changes. The performances of current fNIRS-based BCIs are much inferior to EEG [Coyle 2007]. However, unlike the MEG and fMRI, this technique is portable and inexpensive and therefore might be an alternative in the future.

2.2 Neural mechanisms used in EEG-based BCI

There are currently three major neuromechanisms to control EEG-based BCIs: control of sensorimotor rhythms (SMR), visual evoked potentials (VEP) and P300 event-related potentials (ERP). In the first approach, users learn to control sensorimotor rhythms by performing mental tasks, such as motor imagination (e.g., imagination of hand movement). In the second approach, BCI is controlled by a neurophysiologic response to some type of visual stimulus. The third approach relies on the response to an infrequent event occurring randomly within frequent standard events (oddball paradigm). These events can have different modalities such as visual, auditive or tactile. These three approaches were all experimentally tested at ISR during this thesis. A brief description is given below and a comparison of the methods is provided in Table 2.1.

Sensorimotor rhythms Sensorimotor rhythms are recorded over the sensorimotor cortex (see Fig. 2.1). The main neurophysiologic features are related to the desynchronization and synchronization (ERD/ERS) of sensorimotor rhythms μ (8-12 Hz) and β (18-25 Hz) [Pfurtscheller 1998, Fabiani 2004]. These rhythms typically decrease (ERD) during a motor movement or on its preparation, and increase (ERS) during motor relaxation (see Fig. 2.2a)). Individuals can learn to control SMR by performing sensorimotor tasks. By exploring the topographic map of the motor cortex, distinct regions can be activated by selecting specific tasks for imagination, for instance, imagination of left hand vs. right hand or foot vs. hand [Pfurtscheller 1998], or even with different types of movements for the same limb (e.g., rotation vs. flexion of wrist) [Vuckovic 2008]. Due to the low spatial resolution of EEG, the number of discriminative brain patterns is usually limited to 3-4 [Naeem 2006]. Individuals must train to learn to control their rhythms, which can take several days, weeks or months. However, some people are unable to achieve this control. Visual feedback acts as reinforcement to increase the learning rate.

During SMR control, a motor related potential (MRP) occurs reflecting the prepa-

ration of a movement. This potential is characterized by a slow negative shift, designated by readiness potential (RP) or Bereitschaftspotential, that occurs before the onset of the movement [Cui 1999] (see Fig. 2.2b)). This MRP can be used by itself to control a BCI or in combination with SMR [Blankertz 2006].

Visual evoked potentials A visual evoked potential is an exogenous response to a visual stimulus. It is considered an early component of an event related potential (ERP), i.e., it occurrs within the 100 ms after the onset of the stimulus (e.g., component N100 in Fig. 2.2d)). The response, recorded at visual cortex (see Fig. 2.1), depends mainly on the physical attributes of the stimulus. To control a VEP-based BCI, the user has to direct the gaze to the target stimulus. Therefore, these BCIs are considered dependent BCIs. Compared to SMR and P300, this type of interface usually presents higher communication rates, however, if the user can deftly move the head or eyes, other standard alternatives can be used instead (e.g. head-tracker or eye-tracker). Initial tests with more than two decades [Sutter 1992] showed that subjects were able to control a speller device based on VEPs produced by 64 brief visual stimuli. Currently, most of the approaches are based on steady state visual evoked potentials (SSVEP), which are potentials evoked at the visual cortex as response to stimuli flickering at rates above 6 Hz [Gao 2003] (see Fig. 2.2c)). Recently, other VEPs have been researched for BCI, namely the motion-onset VEP (mVEP), a visual motion response of the higher visual system [Liu 2010]. Although VEP-based BCIs are considered dependent BCIs, in [Kelly 2005] covert attention was used for online SSVEP BCI control, and recently in [Allison 2008] some evidences were reached showing that overlapping stimuli can evoke changes in SSVEP activity sufficient to control a BCI, not requiring gaze shifting. However, conclusive results require further validation.

P300 P300 is an endogenous positive polarity component of an ERP that is elicited by a relevant (target) event in an oddball paradigm (2-state condition where an expected but infrequent event occurs among standard events (non-target)). The stimuli can be a visual, auditory, tactile or somatosensory [Sanei 2007]. P300 waveform is typically characterized by its peak amplitude, latency, and spatial distribution. It occurs typically at a latency of approximately 300 ms after the onset of the stimulus and has a greater centroparietal distribution [Comerchero 1999, Friedman 2001, Donchin 2000] (see Fig. 2.2d)). It is also sometimes referred has P3b in contrast to P3a. This one has a greater frontocentral distri-

bution and is produced by an infrequent and novel (not expected) nontarget event (3-state oddball paradigm) [Comerchero 1999, Friedman 2001]. While VEPs are automatically evoked by external stimuli and depend mainly on their physical attributes, late ERPs such as P300 depend on volitional intent of the user, strongly depending on internal cognitive processes. P300 ERPs vary according to stimulus modality, stimulus probability, inter-stimulus interval, task-relevance, decision making, selective attention, expectancies and relative perceptual distinctiveness among stimuli [Polich 2007, Rohrbaugh 1974, Polich 1996, Comerchero 1999, Heinrich 2008]. An ERP elicited by an oddball paradigm exhibits other components besides the P300 (Fig. 2.2d)). Early components such as the N100 VEP, are automatic responses to sensory input (visual cortex) and subsequent deflections such as P200 reflect early stimulus evaluation and feature detection [Luck 1994]. The N200 component reflects attention, classification and spatial search tasks related to deviant stimuli. P300 reflects broad recognition and memory-updating processes [Patel 2005]. To elicit a P300 potential, users are not required to gaze the target stimuli. Therefore P300-based BCIs are considered independent. However, to perceive the occurrence of the event, users may need to slightly direct gaze to a target stimulus. For example, symbols that appear at the visual periphery may be unattended by the user. Thus, users may have to move their eyes to re-locate the target in the foveal region. These and other related issues will be addressed in more detail in chapters 6 and 7. The term P300-based BCI has also been questioned because besides the P300 component, other components (e.g., N200) have a contribution to BCI performance (see chapter 6). Treder et al. (2010) have suggested to use the name ERP-based BCI instead of P300-based BCI.

See table 2.1 for a summary of generic advantages and disadvantages of the above three neuromechanisms. In addition to the aforementioned approaches, slow cortical potentials (SCP) have been used in BCI until recently [Birbaumer 2000, Hinterberger 2004]. SCPs are very slow positive and negative shifts (up to 10 s) of EEG, usually associated with cognitive tasks. The investigation of these potentials for BCI has been gradually abandoned due to its very low transfer rate and because of the user intensive training. However, this approach has been used for neurotherapy (e.g., control of epyleptic seizures and attention deficit disorders) [Birbaumer 2006].

Table 2.1: Summary and comparison of main neuromechanisms used in EEG-based BCI. (+) stand for advantage and (-) for disadvantage.

	Training	Task difficulty	Task implementation	Type of control	Dependent/ Independent
SMR	Hours to months and some people cannot learn to control (-)	High concentration (-)	No external stimuli required (+)	continuous control (+)	independent (+)
SSVEP	Not required (+) but some people do not evoke	No special requirement (+), but stimuli cause high fatigue (-)	External stimuli required (-) and must be synchronized with refresh rate of display or requirement of external device (-)	discrete control (-)	dependent (-)
P300	Not required (+)	High attention (-), stimuli cause moderate fatigue	External stimuli required (-) and strict synchronism between events and recorded data (-)	discrete control (-)	independent (+)

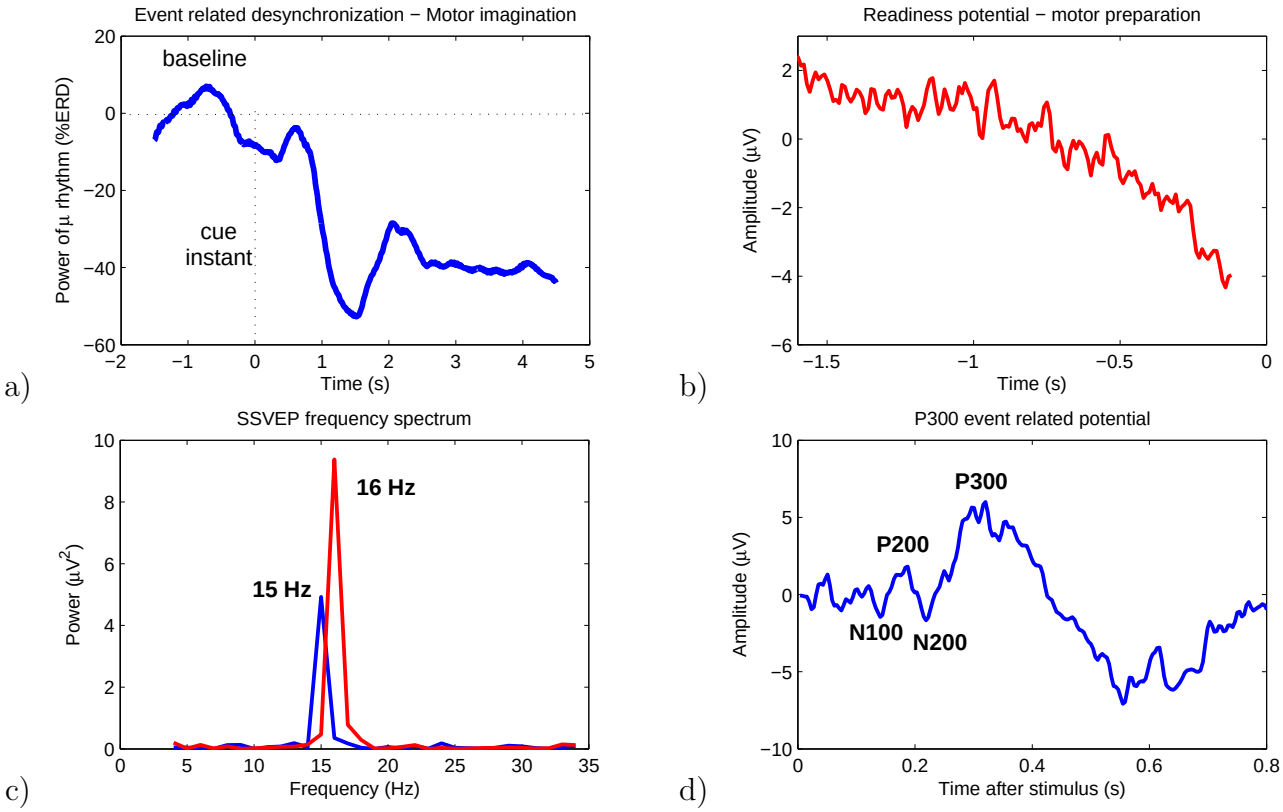


Figure 2.2: a) Event related desynchronization during a motor imagination task (μ rhythm). Cue instant indicates the instant when the user starts the SMR control [Pires 2008a]; b) Readiness potential (Bereitschaftspotential) measured at channel C3. It is a slow negative shift occurring before a finger movement (keyboard press) at instant 0 s [Pires 2007]; c) Frequency spectrum of SSVEP responses to 15Hz and 16Hz flicker stimuli, measured at channel Oz [Ribeiro 2009]; d) Example of a typical P300 event related potential elicited by the LSC paradigm (see chapter 6), measured at channel Cz.

2.3 BCI applications

In the last decade, the research work in EEG-based BCI significantly increased. Every year, dozens of interesting works are published. We present here only some representative works regarding main BCI applications, milestones and trends. See [Blankertz 2010, Mak 2009, Birbaumer 2006] for more extensive surveys. Our review is limited to BCI works based on EEG.

BCI applications can broadly be divided in: 1) communication; 2) control of assistive robotic devices; 3) neurorehabilitation; 4) neurotherapy; and 4) non-medical applications.

Communication - The restoration of communication for people who are 'locked-in' or suffer from severe motor disabilities has been the main topic in BCI research. Several speller

devices have been developed allowing disabled persons to sequentially select letters of the alphabet.

The first speller device was the RC visual P300-speller, introduced by Farwell and Donchin [Farwell 1988]. The paradigm is based on a 6×6 matrix of letters where rows and columns flash randomly (see chapter 3). Since then, one of the most interesting evolutions of the RC speller was the checkerboard speller which extends the rows/columns concept [Townsend 2010]. P300-based communication systems have been successfully used by individuals with neuromotor diseases and motor disabilities resulting from stroke and spinal cord injuries [Kubler 2008]. Typical average rates of P300-spellers achieved by able-bodied individuals are in the range of $20 - 25$ bpm. Bit rates achieved by disabled persons strongly depend on the level of disability. Very recently, a commercial P300 speller for general audience has been made available by g.tec® company [gtec 2009]. ERP-BCIs based on auditory stimuli are being currently researched as an alternative for patients in advanced stages of ALS who show severe eye restrictions (e.g., complete 'locked-in' sate) or individuals with visual impairments [Furdea 2009, Guo 2010, Hohne 2011].

The SSVEP neuromechanism has also been used as a speller device. Several approaches are described in [Wang 2006, Cecotti 2010] showing average bit rates in the range of $30-45$ bpm for able-bodied individuals (no results are reported with disabled patients). These bit rates are achieved by systems where users have to gaze the targets (dependent BCIs).

SMR have been used to control a cursor on a screen with movements in 1D, 2D and 3D [Wolpaw 2004, Neuper 2006, McFarland 2008a, McFarland 2010]. After some training, users learn to control SMR through motor imagination tasks, acquiring the skills to select letters or icons associated to specific control tasks. Reported studies show successful results with patients in advanced stages of ALS, despite the low transfer rates (less than 2 symbols/min). Besides visual feedback, auditory feedback can be provided when individuals have visual impairments [Nijboer 2008].

The regulation of SCP has been used in one of the earliest speller systems, the thought translation device (TTD) developed in Birbaumer's lab [Birbaumer 2000] [Kübler 2001]. Users control a cursor that selects letters by performing binary decisions. Despite the successful results reported with ALS individuals, this neuromechanism offers very low transfer rates and requires long-term training [Birbaumer 2006].

Wheelchair steering - Paralyzed individuals bounded to a wheelchair and unable to control standard interfaces may benefit from BCI to restore their mobility. However, this implementation faces many challenges. First, the BCI has to function asynchronously, i.e., the user decides when to send a command and not the system itself. Therefore, a 'non-control' state should be detected additionally to steering commands. Second, since the number

of decoded symbols per minute is very low, usually less than 10 commands per minute, wheelchair manoeuvres have to be provided by a semi-autonomous navigation system that receives sparse commands from user. Millan performed the first experiments in this field, controlling a mini-mobile robot [Millan 2004]. Since then, several experiments have been performed toward brain-actuated wheelchairs, using SMR [Tanaka 2005, Galán 2008] and P300 [Iturrate 2009, Rebsamen 2010], addressing the asynchronous operation issue. The success of these systems strongly depend on mobile robot advances.

Environmental control - In the actual societies, the concept of ambient assisted living (AAL) has an increasing social importance. The control of simple environmental domestic devices through BCI can substantially increase the level of independence of home-bounded individuals who suffer from severe motor disabilities. In [Gao 2003], SSVEP controls an infrared remote-controller that actuates multiple external devices, such as TV, video tape recorders, or air-conditioners. A complete AAL system was designed in [Cincotti 2008]. The home environment was tested by several motor disabled patients who could control external devices such as neon lights and bulbs, TV, motorized bed, acoustic alarm, doors and telephone. These devices were commanded by selecting icons in a computer screen through modulation of SMRs. A P300 neuromechanism was also used to control an apartment in a 3D virtual environment [Bayliss 2003].

Neuroprosthetics - Neuromotor prostheses can allow tetraplegic individuals or amputees to restore some movement abilities. Due to the low transfer rates provided by BCI, this is a very challenging research area. Yet, successful invasive and non-invasive applications have been reached [McFarland 2008b]. In [Pfurtscheller 2000], a tetraplegic patient using an SMR-based BCI, learned to control an electrical driven hand orthosis, restoring its hand grasp function. Functional electric stimulation (FES) of hand muscles controlled by SMR was also applied to restore hand grasping [Pfurtscheller 2005]. Tasks such as the control of a 3D cursor or the control of a robotic arm are more complex. These tasks usually follow two approaches: 'goal selection' or 'kinematic control' [McFarland 2008b]. The second approach is very demanding, since it requires the BCI to provide continuously 3 control signals (one per dimension). This approach has been used mostly in intracortical BCIs with tests in monkeys [Taylor 2002], while the less demanding 'goal selection' has been used mostly in non-invasive BCIs [McFarland 2008b].

Neurorehabilitation - It was shown in [Daly 2008] that neurotherapy based on BCI can be used in parallel with standard motor therapies to help stroke patients to recover injured neuromuscular functions. The underlying idea is that BCI might be used to induce brain plasticity activity. According to [Daly 2008], learning to regain motor control can be performed following two strategies: 1) by providing visual feedback of EEG activity, users

are asked to regulate SMR activity [Daly 2006]; or 2) by translating the EEG activity to activate a device that assists movement [Daly 2009]. Learning strategies have to be used adequately, otherwise abnormal motor function can be even exacerbated.

Neurotherapy in behavioural disorders - Neurofeedback, i.e., the real-time feedback of brain activity, has been used for some time to ameliorate ADHD in children [Monastra 2005, Gevensleben 2009]. For instance, in ADHD, EEG rhythms related to attention can be translated into game commands, e.g., control the altitude of a plain or similar tasks [Wang 2011]. In the example, a child who is asked to maintain the altitude of the plain above a threshold level, is actually controlling its attention level. These systems are today an effective alternative to pharmacologic treatments. Same neurofeedback principles can be used for therapy in autism, dyslexia, tourettes, to control seizure frequency in people with epilepsy, or for neuroreahabilitation as described above.

Non-medical - BCI techniques can be applied for non-medical purposes for healthy people. EEG monitoring for car accident prevention, monitoring of performance capabilities, workload evaluation and cognition evaluation, are examples of applications that can have positive outcomes in our daily lives [Blankertz 2010]. Game-control based on BCI is currently used in research context [Pineda 2003, Reuderink 2008, Tangermann 2008, Finke 2009], however it will be for sure present in our daily lives, in a near future.

2.4 Conclusion

Although invasive BCIs may potentiate higher transfer rates and a broader range of real-world applications, several researchers state that current non-invasive BCIs can provide performances comparable to that achieved by invasive methods [Daly 2008] [McFarland 2010]. The main limitative factor for the effective use of BCI in real-world scenarios is its low transfer rate. Moreover, the control of robotic and environmental devices require an asynchronous participant controlled operation, not available in most of the systems. From the three neuromechanisms used in non-invasive BCIs (P300, SSVEP and SMR), SSVEP presents the higher transfer rates. However, because SSVEP is gaze dependent, this option has been discarded from most of the clinical studies. P300-based BCIs show transfer rates substantially higher than that achieved with SMR-based BCIs in clinical studies with severe motor disable patients [Nijboer 2010]. However, when a continuous control is required, stimuli-based BCIs may not be the most suitable, since they only provide discrete commands in time. The choice of a neuromechanism to control a BCI should take several aspects into consideration: user's limitations, user comfort,

specific requirements and usability of the application.

Chapter 3

P300-based BCI: Material and methods

Contents

"Information is the resolution of uncertainty."
Claude E. Shannon

This chapter describes the complete design of a P300-based BCI system. It provides the background and materials necessary to the remaining chapters. Although this chapter

focuses the standard Row-Column speller (RC), the underlying architecture is common to any other P300-based BCI. All aspects associated with a P300-based BCI are explained, namely, paradigm/protocol, terminology, synchronization and performance metrics. The software framework that was developed for the implementation of the P300 systems is succinctly described.

Additionally, a description of the disorders of disabled individuals who took part in experiments is provided.

3.1 Standard Row-Column P300 Speller

The RC speller, introduced by Farwell and Donchin in the eighties [Farwell 1988], was the first BCI system based on P300. Nowadays, this paradigm is still the P300 system mostly used for spelling and remains a benchmark for system comparison. Our implementation follows the original RC paradigm, despite some slight modifications, such as the set of symbols and the flashing presentation.

3.1.1 Paradigm description

The layout of the RC speller paradigm is a 6×6 matrix that comprises the alphabet letters and other useful symbols such as the 'spc' and 'del', which are used respectively to separate the words and to delete misspelled letters (see a screenshot in Fig. 3.1a)). The rows and columns of the matrix flash randomly according to an oddball paradigm. The user mentally selects a letter. The target event corresponds to the row or column that includes the symbol mentally selected (relevant stimuli), while all other rows and columns correspond to non-target (standard) events (see screenshots of successive events in Fig. 3.1b)). Target events are supposed to elicit a P300 component. The detected symbol is obtained from the combination of the target row and target column. From the 12 events (6 columns and 6 rows), two are relevant targets, thus target probability is (1:6). The rows and columns are flashed (highlighted) by changing the color from gray to white. To increase flash perception the size of the symbols are slightly increased by $\approx 15\%$ when highlighted. The interval between the onset of two consecutive stimuli (stimulus onset asynchrony, SOA) and the inter-stimulus interval (ISI) are both adjustable (Fig. 3.1)c)). In most of the experiments, SOA and ISI were respectively set to 200 ms and 100 ms. A sub-trial consists of a round of 12 events as shown in Fig. 3.1c). Since the SNR of the EEG signal is very low, it is usually necessary to combine several sub-trials to correctly detect the targets. The number of sub-trials is settled before the

online operation, however some approaches can be used to automatically adapt/adjust the number of event repetitions according to user performance in the course of the online experiment [Lenhardt 2008, Jin 2011b]. The recorded EEG must be synchronized with the onset of each event, and labeled with the code of the respective event. The data segment associated to each event is called an epoch which has a duration of 1 second in our experiments (Fig. 3.1c)). An interval between trials (inter-trial interval, ITI) is required to allow the user to switch the attention focus for a new symbol mentally selected.

The terms symbol/character/letter are sometimes interchangeably used always with the same meaning.

3.1.2 Classification architecture

Before the online operation, a P300-based system requires a calibration phase, from which EEG is recorded and labeled according to the event types (target and non-target). During this phase, the user mentally selects pre-defined letters, which are sequentially provided at the top of the screen. The labeled data (ground truth) are used to train, in offline, the feature extractors, the feature selectors and the binary classifiers models, as illustrated in Fig. 3.2-Top and are then applied online (Fig. 3.2-Bottom). A P300-based system comprises two types of classification. A binary (two-class) classification and a C-class classification for character recognition (the number of classes is the same as the number of symbols). These two classifiers are applied sequentially:

Binary classifier: taking the models obtained for target and non-target epochs, the binary classifier is applied to each data segment (epoch) associated to an event, and detects whether or not it contains a P300 component.

C-class classifier: the outputs of the previous binary classifications are combined to find the most likely target events. The output class and respective score of each event is used to find the character mentally selected. This decision module plays an important role in the P300 system and may significantly improve the online accuracy. The implementation also depends on how the sub-trials are combined. An approach is detailed in section 5.4.4.

In particular for the RC speller, the number of symbols (classes) is given by $(N_{ev}/2)^2$ where N_{ev} is the number of events. For a 6×6 matrix, $N_{ev} = 12$ and thus the symbol recognition is a 36-class classifier. It should be stressed out that random level accuracy is 2.7% for $N_s = 36$ symbols. After the C-class classifier, a word prediction can further be applied

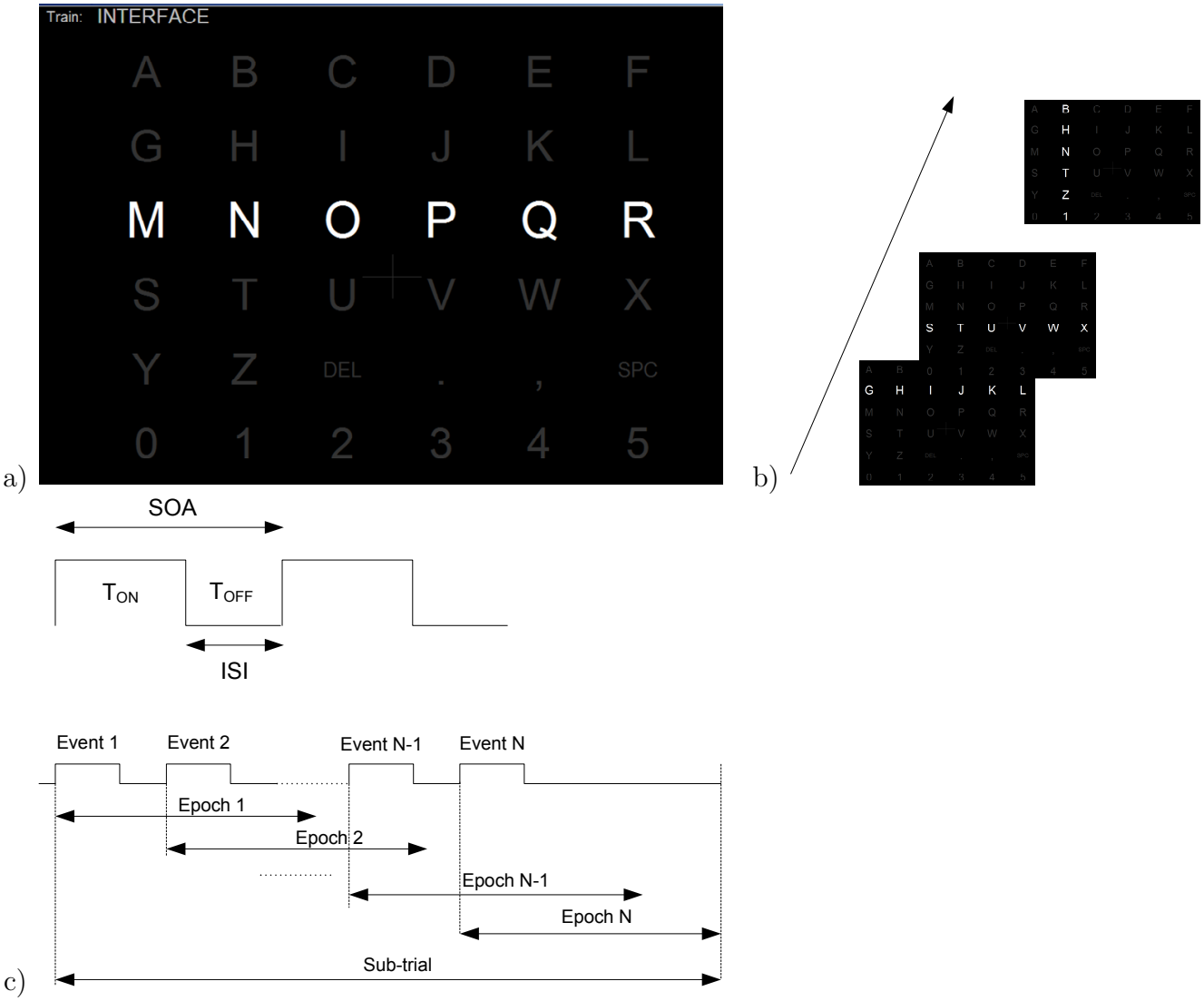


Figure 3.1: a) Screenshot of 6×6 matrix speller paradigm; b) Screenshots of successive events; c) Top: temporal diagram defining SOA and ISI. Bottom: temporal diagram showing epochs extracted from the continuous EEG data stream.

driven by dictionaries (intelligent writing) [Ahi 2011]. Thus, usually the accuracies are provided for the binary classifier and for the character recognition classifier.

3.2 Performance metrics

To assess the performance of a BCI system, several evaluation criteria have already been proposed (see a survey in [Schlögl 2007]). A BCI system encompasses many parameters related to its underlying paradigm, which includes the amount of encoded information,

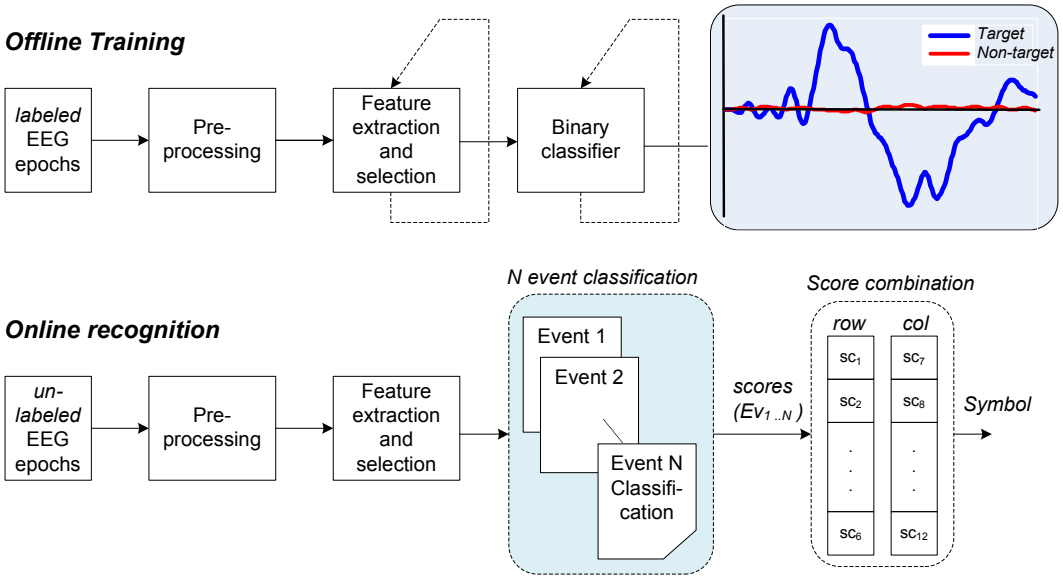


Figure 3.2: Classification architecture of a P300-based BCI system. Top: offline training to obtain feature extractor models, selected features, and classifier models. Bottom: for the online recognition, the binary classifiers are applied to each event, and then their output scores are combined to identify the mentally selected symbol.

the required time for selecting a symbol and its detection accuracy. The ultimate goal of a BCI is to provide an effective communication channel. Thus, the gold standards to evaluate a BCI performance should be the online accuracy and the online bit rates. Only these evaluation criteria can attest the effective application of BCI in real world scenarios. Offline classification accuracy (or error rate) is a powerful measure to train classifiers, but it only gives an estimation of the potential online performance [Townsend 2010]. While some research works show impressive offline results, they lack of rigor for two reasons: 1) the results concern the binary classification only, omitting the $C-$Classification part; 2) the offline binary classification use inappropriate accuracy metrics. This section describes and discusses the online metrics and the offline evaluation criteria used in this thesis.

3.2.1 Bit rate metrics

Wolpaw *et. al* introduced a metric to compute the information transfer rate (ITR) of a BCI system [Wolpaw 2000]. The formula can be obtained from Shannon's [Shannon 1948] theory by modeling the BCI system as a noisy communication channel (see Fig.3.3). The average mutual information, $I(X;Y)$, between the intention of the user and the detection

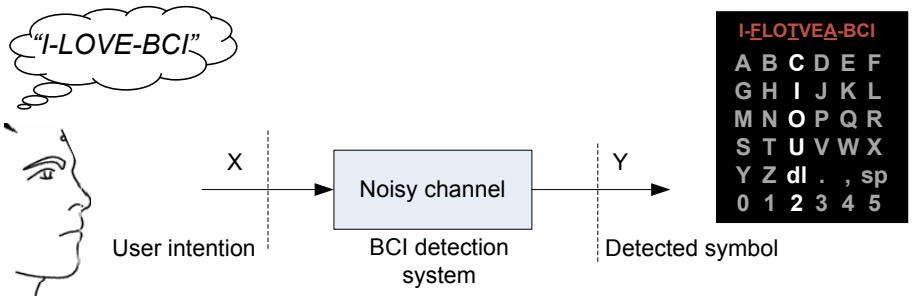


Figure 3.3: BCI system modeled as a communication channel. Simulation of an online experiment with 3 errors in 10 characters (70% accuracy).

made by the BCI system is given by

$$I(X;Y) = H(X) - H(X|Y) \tag{3.1}$$

where $H(X)$ is the source entropy and $H(X|Y)$ is the information lost in the noisy channel, i.e., it represents the classification error rate of the BCI system. Assuming a BCI with N_s possible choices (number of symbols) which are equiprobable, and an online classification accuracy of P_{ac}, then $I(X;Y) \equiv B$, measured in bits/symbol, is given by (see Appendix A.1 for a derivation of this formula)

$$B = \log_2(N_s) + P_{ac}\log_2(P_{ac}) + (1-P_{ac})\log_2\frac{(1-P_{ac})}{(N_s-1)}. \tag{3.2}$$

Taking the rate of possible selections per minute r_s (symbols per min, SPM) then the ITR is expressed by

$$ITR = r_s B. \tag{3.3}$$

The r_s rate is obtained from

$$r_s = \frac{60}{N_{rep} \times (N_{ev} \times SOA) + ITI}. \tag{3.4}$$

This metric is currently widely used by the BCI community as a benchmark metric for performance comparison, particularly in P300-based BCIs. The ITR reflects simultaneously the accuracy, the number of symbols per minute and the amount of encoded information. However, the use of this metric for the assessment of a BCI should always be accompanied with the online accuracy. Metric (3.3) can be fallacious because low levels of accuracy may provide reasonable bit rates while at the same time be unacceptable for an effective communication [Sellers 2006a]. For example, in Fig. 3.3, a simulation of an

online experiment spelling "I-LOVE-BCI" with 70% accuracy, shows that the detected sentence is nearly unintelligible. Nevertheless, for this classification accuracy the ITR is more than 12.5 bpm (see graph in Fig. 3.4).

Actually, ITR does not take into account the correction of errors and therefore it can not provide a realistic metric of the overall BCI performance. Several authors [Townsend 2010, Dal Seno 2010] realized that ITR is not the most suitable metric and argue that a more practical and generic metric is needed to assess and compare BCIs. A metric that takes into account the correction of misspelled letters is more suitable because it clearly expresses the effect of online classification accuracy. The time needed to correctly spell a letter can be obtained by computing the average number of tries, N_r until the letter is correctly spelled. Every time an error occurs, two additional selections are required (one for deleting and one for re-spell). This can be seen as a communication protocol for error recovery controlled by the user. The N_r value is computed following an approach similar to that found in [Dal Seno 2010] and [Townsend 2010]. Assuming that user wants to correctly spell N symbols, then, taking the probability of failure (error p_e) and associated overhead, the number of required selections is

$$N + 2(N)p_e + 2(2Np_e)p_e + 2(2(2Np_e)p_e)p_e + \cdots = N\sum_{i=0}^{\infty}(2p_e)^i. \tag{3.5}$$

The series converge to

$$\frac{N}{1-2p_e} \tag{3.6}$$

as long as $p_e < 0.5$. Otherwise, the required number of selection tends to infinite, i.e., the user will never be able to correctly spell a sentence. Rearranging the above equation, in terms of classification accuracy ($P_{ac} = 1 - p_e$), instead of the error rate, we obtain the number of tries, N_r from

$$N_r = \frac{\frac{N}{1-2(1-P_{ac})}}{N} = \frac{1}{1-2(1-P_{ac})}. \tag{3.7}$$

The practical bit rate (PBR) is then obtained from

$$PBR = \frac{r_s}{N_r}\log_2 N_s, \tag{3.8}$$

where $\log_2 N_s$ is the source entropy, $H(X)$. A comparison between the ITR and PBR is shown in Fig. 3.4 taking respectively (3.3) and (3.8). The simulation was made for a

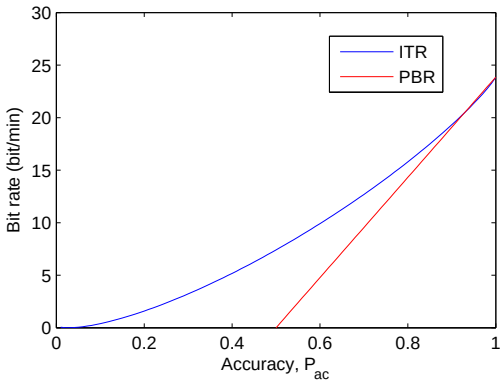


Figure 3.4: Bit rate comparison between ITR (3.3) and PBR (3.8) varying according to classification accuracy. The simulation was made for a RC speller, 36 choice, $N_{rep} = 5$, $ISI = 200$ ms and $ITI = 0$.

6×6 matrix (N_s=36 symbols), with $N_{rep} = 5$, $ISI = 200$ ms and omitting the ITI. For example, for a 50% classification accuracy, the PBR is 0 bpm, while the ITR is 7.40 bpm. Despite this significative ITR value, user would never be able to correctly spell a sentence. In the other hand, assuming no correction, spelling is not intelligible below certain levels of classification accuracies. Approaching $P_{ac} = 100\%$, the values converge as expected.

It is further required to present a suitable formula to measure the online accuracy P_{ac}. Since it is possible to correct a symbol, and the correction ('del' symbol) is itself a selection, both spelled and correction symbols must be taken into account, thus

$$P_{ac} = 1 - \frac{N_e}{N_c + N_{ce}} \tag{3.9}$$

where N_e is the number of misspelled characters/symbols, N_c is the number of characters of the sentence and N_{ce} is the number of corrected errors with 'del'.

Section 7.2.2 in chapter 7 extends the practical bit rate metric (3.8) to a novel 2-level speller (GIBS).

3.2.2 Offline evaluation

3.2.2.1 Accuracy

The offline binary classification accuracy can provide an estimate of the online BCI performance. In P300-BCI, this measurement has to be carefully analyzed, because target and non-target classes are highly unbalanced. This occurs because the probability of one class is much smaller than the other one. For that reason, the performance measures have

to be selected appropriately. Considering the information of true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN), the common classification accuracy is given by

$$P_a = \frac{TP + TN}{TP + FP + FN + TN}. \tag{3.10}$$

Because of the unbalanced classes, this measure is not suitable as we shall see. Suppose a target and non-target probability of 1/10 and 9/10, respectively. If the classifier classifies all epochs as non-target, then the classification accuracy is 90%, despite it failed all target epochs. Obviously, this measure is a poor indicator for online performance. This issue is sometimes disregarded in some works. To deal with unbalanced classes, we propose a balanced accuracy measure, which treats both classes with equal importance, that is

$$P_{bal} = \frac{TPR + TNR}{2} = 0.5\frac{TP}{TP + FN} + 0.5\frac{TN}{TN + FP} \tag{3.11}$$

where TPR and TNR denote respectively true positive rate and true negative rate. Equivalently, the balanced error is given by

$$e_{bal} = \frac{FPR + FNR}{2} \tag{3.12}$$

where FPR and FNR denote respectively false positive rate and false negative rate. If not stated otherwise, these two measures will be the standards to evaluate the offline performance in the remaining chapters, and they will be sometimes simply referred as classification accuracy and error rate. In a clinical case-study presented in appendix D, the sensitivity and specificity are also used as performance measures

$$Sensivity = \frac{TP}{TP + FN} \qquad Specificity = \frac{TN}{TN + FP}. \tag{3.13}$$

As a remark, note that the balanced accuracy (3.11) embodies simultaneously specificity and sensitivity (3.13).

3.2.2.2 SNR

The SNR measure provides important and reliable performance information. From a neurophysiological perspective, it assesses the 'quality' of the ERP, by measuring the inter-trial variability, and it quantifies the improvements achieved by the application of signal processing techniques. From a classification perspective, our own observations clearly indicate a direct relationship between SNR and classification performance (see chapter

5). This observation makes SNR a good measure to train or tune signal processing and classification algorithms. Considering a set of K target epochs $\mathbf{y}$, the SNR is estimated according to [Lemm 2006]

$$\begin{aligned} \mathrm{SNR}(\mathbf{y}) &= 10 \log \frac{\mathrm{Var}_t(\mathrm{E}_k[\mathbf{y}])}{\mathrm{E}_k[\mathrm{Var}_t(\mathbf{y} - \mathrm{E}_k[\mathbf{y}])]} \\ &= 10 \log \frac{\mathrm{Var}_t(\bar{\mathbf{y}})}{\mathrm{E}_k[\mathrm{Var}_t(\mathbf{y} - \bar{\mathbf{y}})]} \end{aligned} \tag{3.14}$$

where Var_t is the temporal variance of the ERP signal and $\mathrm{E}[\cdot]$ is the expected value estimated from all K epochs.

3.3 Experimental setup

Experiments consisted on calibration and online feedback sessions. During the calibration sessions, the user of the BCI system is asked to perform a task, according to a visual paradigm. The recorded data are labeled with class information (ground truth). These data are used to train the classification models and then applied in online sessions. The classification models are specific to each subject.

The experimental setup is presented in Fig. 3.5. An EEG amplifier and acquisition system (gtec® product [g.tec]) with 16 channels (scalable to 64 channels) was used for data recording. The electrodes location for a 65 channel g.tec cap is shown in Fig. 3.6. Depending on the experiments, EOG was also recorded in addition to EEG.

3.3.1 Software framework

The overall BCI system, including signal processing, classification and visual paradigms, was implemented in a Simulink® framework [Mathworks 2010a]. A P300-BCI system requires a strict synchronization between recorded data and events, and encompasses several other issues, as explained in section 3.1. Signal acquisition and real-time control loop is made at 256 Hz rate. Recorded data is synchronized with stimuli events; the delay between event generation and stimuli presentation is minimum and constant; signal processing and classification algorithms are computed in real-time; and the control output is immediately visualized. To ensure a real-time operation, a 'Highspeed' Simulink driver provided by g.tec was used. A generic Simulink model of the BCI system is shown in Fig. 3.7. Respective blocks are succinctly described as follows:

Figure 3.5: Picture of the experimental setup at ISR.

'Acquisition driver' block: the gUSBamp driver provides a hard real-time clock through an hardware interrupt that drives the whole Simulink model. This ensures the model and respective blocks to update every time sampling [Guger 2009]. It reads the EEG flow sent from gUSBamp via a USB connection. The gUSBamp block allows also the configuration of some hardware parameters such as selection of sampling rate, pre-set filters, monopolar and bipolar montages, etc.;

'Preprocessing' block: performs temporal filtering or other basic preprocessing;

'Event synchronization and classification' block: implements the event generation and synchronization, data buffering, epoch segmentation, and implements the algorithms for EEG signal processing and classification;

'Visual paradigm' block: implements the visual paradigms. The following were implemented: RC speller, LSC speller, GIBS and Arrow paradigm;

Signal processing, classification and visual paradigms were programmed recurring to Simulink S-functions [Mathworks 2010b]. An S-function uses a special syntax called the S-function API (application program interface) that enables the interaction with the Simulink engine. Temporal events generated in the 'Event synchronization & classification' block are fully parameterized (e.g. ISI and ITI). The classification algorithms are directly embedded in the 'Event synchronization & classification' block after they have been trained offline in Matlab®. This means that no significant changes have to be made between offline training and online implementation. At the end of each trial, the 'Event synchronization & classification' block sends the control output to the 'Visual paradigm' block. The visual paradigms RC speller, LSC speller, GIBS and Arrow paradigm, are

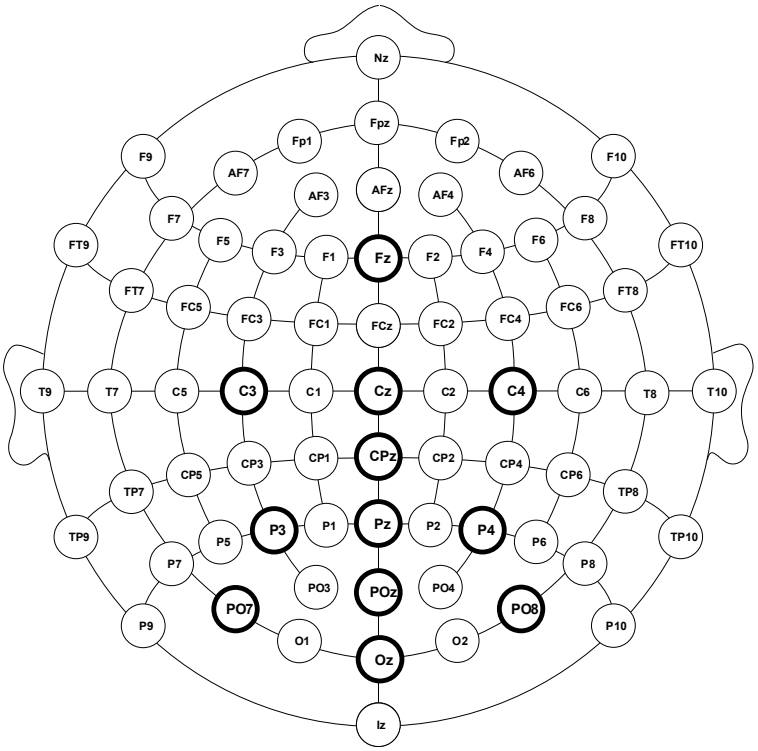


Figure 3.6: Electrode locations, according to the international 10-20 extended system, of the 65 channel cap used in our experiments (g.EEGcap). Bold circles indicate the channels used in most of the P300 experiments.

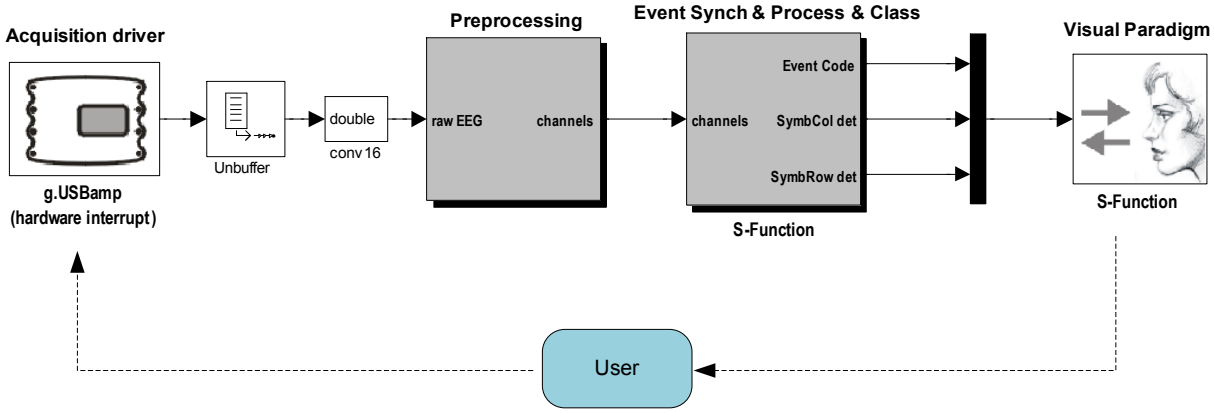


Figure 3.7: Main blocks of a general Simulink implementation of the BCI system, and representation of User in the loop.

based on text and simple graphics and thus their implementation in S-functions was not an issue. However, in case of a Tetris game paradigm, presented in appendix C, the graphical part of Tetris required an enhanced artwork. Tetris graphics were implemented externally to Simulink based on the Tcl/Tk high level language. The data communication between Simulink and the external application relied on a shared memory structure (see [Pires 2011a] and [Casaleiro 2010] fore more details). This general Matlab/Simulink framework provided flexible and efficient prototyping, allowing a fast implementation of

signal processing and classification algorithms in Matlab and C language, and an easy parametrization and reconfiguration of the visual paradigms.

3.4 Motor disabilities of participants

Fourteen subjects with motor impairments participated in the experiments described in this thesis, namely, 7 subjects with amyotrophic lateral sclerosis (ALS), 5 subjects with cerebral palsy, 1 subject with Duchenne muscular dystrophy (DMD), and 1 subject with a spinal-cord injury (SCI). All participants gave informed consent, approved by the Cerebral Palsy Association of Coimbra (APCC) and by the Hospitals of the University of Coimbra (HUC), to participate in the study. The sessions with CP, DMD and SCI subjects took place at the facilities of APCC, and the sessions with the ALS subjects took place at the facilities of HUC. This section aims to briefly describe these motor disorders and to define related terminology, providing the necessary background to the following chapters.

3.4.1 Amyotrophic lateral sclerosis

Amyotrophic lateral sclerosis also referred as Lou Gehrig's disease is a motor neuron disease (MND). It is a neurodegenerative disease that results in progressive degeneration of lower and upper motor neurons in the brain and the spinal cord (central nervous system) [Ioannis 1998]. Motor neurons are responsible to send the commands to voluntary muscular contraction. Thus, with the loss of motor neurons, muscles are unable to function, leading to muscle weakness and atrophy. ALS is classified as either of bulbar-onset, characterized by progressive dysphagia (swallowing difficulty) and dysarthria (poor speech articulation), or of spinal-onset marked by peripheral neurologic features (muscle atrophy, cramps, and fasciculations), and by spasticity (stiff muscles) [Leigh 1994, Fiorentino 2009]. In the advanced stage of the disease, patients can progress to a locked-in state (LIS), with only residual voluntary control, such as eye movement and blinks. In the end-stage of the disease, patients may lose all motor control entering in a complete locked in state (CLIS) [Kubler 2008]. The degree of disability is rated by a revised ALS functional rating scale (ALSFRS-r) where 48 is normal and 0 a complete loss of functionality [Cedarbaum 1999].

3.4.2 Duchenne muscular dystrophy

Duchenne muscular dystrophy is a genetic disorder caused by a single gene mutation. This mutation disrupts the production of a neuromuscular protein, which results in the

gradual breakdown of muscle fiber [Cyrulnik 2008]. Symptoms usually appear in male children before age 6. Children and adolescents with DMD evidence a progressive muscular weakness affecting all voluntary muscles and often exhibit cognitive deficits. Usually by the ages of 12, children lose the ability to walk and are wheelchair dependent. In later stages, DMD affects heart and respiratory muscles, and thus patients require artificial ventilation and other types of support.

3.4.3 Cerebral palsy

Cerebral palsy is a non-progressive neurological disorder resulting from a brain injury that occurs before cerebral development is complete [Krigger 2006]. The etiology can be prenatal, perinatal, or postnatal. It is generically characterized by abnormal movements and posture usually accompanied by dysarthria. The Surveillance of Cerebral Palsy in Europe (SCPE) divides CP into three groups based on the predominant neuromotor abnormality: spastic, dyskinetic, or ataxic, with dyskinesia further differentiated into dystonia and choreoathetosis [Bax 2005]. CP is also classified by topographical pattern of limb involvement, such as diplegia, hemiplegia, or quadriplegia [Jones 2007], however a terminology indicating explicitly the number of limbs affected is now preferred. The CP subtypes are briefly defined: spastic (hypertonia in the muscles that result in stiffness); dyskinetic (involuntary movements, variable muscular tonus); dystonia (hypertonia and involuntary movements); choreoathetosis, combination of chorea (irregular migrating contractions) and athetosis (twisting and writhing); ataxic (poor balance and equilibrium and uncoordinated voluntary movement). Frequently, subjects have associated cognitive impairments. The symptoms and functional levels vary significantly among subjects. In extreme cases, subjects have neither motor control nor communication due to cognitive impairment. Many others are wheelchair bounded but retaining cognitive capabilities sufficient to control adapted interfaces for communication. Other subjects have an almost normal life, despite some motor limitations.

3.4.4 Spinal cord injury

A spinal cord injury (SCI) can be caused by trauma, disease or congenital disorders [NIH 2011]. Depending on the level of injury, damage to the spinal cord can result in paralysis of the muscles used for breathing, paralysis and/or loss of feeling in all or some regions of the trunk, arms, and legs. Tetraplegia, caused by the damage of cervical region (C1 through C7), is one of the most severe conditions in which the individual can

experience a total loss of movement in limbs, torso and neck.

Chapter 4

Classification methods in P300-based BCI

Contents

"No free lunch theorem."
(There is no overall superiority of a particular learning or recognition algorithm)
David Wolpert

This chapter describes state-of-the-art techniques currently used for classification of event related potentials. Concepts and methodological issues related to implementation are also succinctly explained. A classification system is complex and deals with

many issues. Its performance depends on feature extractors, feature selectors and classifiers. Several critical properties, ranging from feature non-stationarity, noise, outliers, dimensionality, size of training sets, learning approach, bias-variance tradeoff, need to be addressed. See a generic review in [Lotte 2007].

This chapter reviews some of the techniques that were tested offline or online in the context of one or more of our proposed P300-based BCI systems, or in benchmark datasets.

4.1 Introduction

EEG has a substantial variability between subjects, within the same subject in the course of different experimental sessions, or even within the same session due to psychophysiological factors (e.g., fatigue). Thus, the performance of signal processing and machine learning techniques can be greatly increased if they are fitted specifically to a subject before an experimental session. Moreover, subject-specific (subject-dependent) models significantly reduce the complexity of classification techniques. However, the main inconvenience of subject-dependent models is that it needs a calibration before the online operation of the system. This calibration involves two parts: 1) EEG recording session, where user performs a task according to pre-defined cues, i.e., EEG data is associated with true labels (ground truth); and 2) using the acquired labeled data, signal processing and machine learning techniques find the features that optimize subject's performance for the task.This supervised approach (training from examples) is the most common in BCI. The time required to perform the calibration depends on the type of classifier being used. For example, some classifiers require a large amount of training data to obtain good generalization to unseen data. Moreover, complex classifiers may require a long time to learn and tune parameters. Obviously, the elimination of the calibration, i.e., a zero calibration-time, would be the ideal scenario. This approach requires a subject-independent model obtained offline from a pool of subjects, and then an online adaptation of the model to the user, through unsupervised learning. Although this issue has been addressed in several offline studies, only a few have tested unsupervised learning online with success. Implementations with only a small performance decay are described in [Lu 2009] for P300-based BCI and [Vidaurre 2011] [Krauledat 2008] for motor imagery based BCI. Similar unsupervised techniques can be applied to adjust user-specific model in the time course of the session. In P300-based BCI contexts, session-independent models are more meaningful than subject-independent models. However, in some contexts other than BCI, such as applications for clinical diagnosis (e.g., *case-study VI* in appendix D),

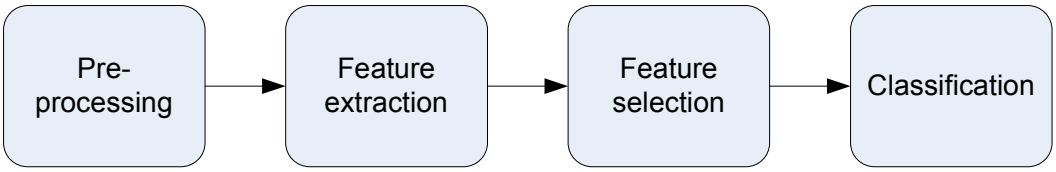


Figure 4.1: Main blocks of a generic classification system.

a subject-independent model is required because of the impossibility of prior calibration.

A classification system is generically composed by four main modules integrating knowledge from signal processing and machine learning areas: 1) preprocessing; 2) feature extraction; 3) feature selection; and 4) classification; (see Fig. 4.1). Regarding the classification system, this thesis contributes on new feature extraction techniques in the spatial domain. Since this is a major theme of this thesis, chapter 5 is devoted entirely to it. Feature selectors and classifiers were used in their standard form, despite some slight rearrangements. Yet, the understanding and correct implementation of these techniques were crucial to achieve the high performance levels of the proposed BCI systems. The rest of this chapter succinctly describes state-of-the-art techniques for feature extraction, feature selection, and classification, in the context of P300-based BCIs, all applied in this thesis. Although the session-dependent issue was not addressed here, calibration time was still of main concern. For all of the proposed paradigms, we limited the calibration time to 5 min for the data recording session and less than 2 minutes to obtain classification models. This calibration time is considerably less than those presented in some reference studies using supervised learning, for example 30 min plus classifier training in [Lenhardt 2008] and 9-12 min plus classifier training in [Townsend 2010].

4.2 Pre-processing and feature extraction

The classification of ERPs is a very challenging problem due to its high inter-trial variability and low SNR. Thus, the main goal of signal processing techniques is to increase the SNR of ERPs and alleviate variability before classification. Basic temporal filtering, such as low-pass, high-pass, band-pass and notch, are usually applied to EEG, to eliminate signals that are outside the frequency range of interest, or to eliminate specific interferences (e.g., powerline). Different EEG derivations, such as common average reference (CAR), bipolar, and Laplacian, act as spatial filters contributing also to an increased SNR [McFarland 1997].

Pre-processing using the above techniques might not be sufficient to provide adequate features for classification, and therefore it is necessary to extract powerful discriminative features. Feature extraction consists of linear or non-linear transformations which aim to decorrelate/decompose features for denoising, maximize the discriminative power of the features, or remove redundancy. At the same time, feature extraction reduces the dimensionality of feature space, by mapping the input data into a lower dimensional space. A low dimensional feature space avoids overfitting to training data, increases algorithm robustness, and provides faster and more effective algorithm computation [Hall 2000]. Moreover, features are more easily interpreted, establishing potential bio-markers with neuro-scientific meaning [Peng 2005].

Feature extraction methods can be applied in time, frequency, time-frequency and spatial domains. [Bashashati 2007] presents an extensive survey on feature extraction methods applied in BCI. In SMR and SSVEP BCIs, neurophysiologic features are related with rhythmic activity, and thus feature extractors are applied mainly in the frequency domain. Spectral estimation based on parametric and non-parametric techniques is applied on raw EEG data. Techniques include direct application of discrete Fourier transform, periodogram based methods (e.g., Whelch), auto-regressive models and wavelets (discrete and continuous wavelet transform (DWT, CWT)). In the other hand, ERPs, such as P300 or MRPs, are components time-locked and phase-locked with events, and therefore they are characterized by their temporal evolution. Some examples of feature extraction techniques applied to P300 include peak picking, area, covariance [Farwell 1988], matched filter [Serby 2005] and DWT [Donchin 2000]. However, in the most common approach, the features are the amplitudes of the time samples at a given interval, previously low-pass filtered and downsampled [Thulasidas 2006, Krusienski 2007, Lenhardt 2008, Rakotomamonjy 2008, Hoffmann 2008a, Zhang 2008]. Downsampling, or optionally averaging across small sub-intervals [Blankertz 2011], reduces the dimensionality of feature space. The feature vector is composed of the concatenation of the features of each recording channel. Considering a spatio-temporal data matrix representing an EEG epoch $\mathbf{X} \in \mathbb{R}^{N \times T}$, where N is the number of channels and T is the number of time samples (or the number of features after feature extraction or downsampling), then the feature vector resulting of the concatenation is given by

$$\mathbf{x} = [x_1^{(1)} x_2^{(1)} \cdots x_T^{(1)} \quad x_1^{(2)} x_2^{(2)} \cdots x_T^{(2)} \quad x_1^{(N)} x_2^{(N)} \cdots x_T^{(N)}] \tag{4.1}$$

which as a dimension $1 \times NT$. This vector is referred here as spatio-temporal feature vector. In the other hand, feature extractors applied in the spatial domain combine

optimally signals recorded at different electrode locations. The projections resulting from spatial filtering are concatenated building a spatial feature vector. The data matrix $\mathbf{X} \in \mathbb{R}^{N \times T}$ is transformed into a new matrix $\mathbf{Y} \in \mathbb{R}^{N' \times T}$, where N' is the number of spatial projections, typically with $N' \ll N$, eventually $N' = 1$. This transformation reduces significantly the feature dimensional space. While this spatial filtering strategy is a common practice in BCIs based on SMR, with proved efficacy, it is not very often in BCIs based on ERPs. In chapter 5, we show that spatial filtering is also effective for ERP classification.

4.3 Feature Selection

To further reduce the dimension of the feature vector, extracted features are selected according to their discriminative power. Although feature extraction and feature selection both reduce feature dimensionality, they are quite different. While the former transforms the features, feature selection only selects the features without transform them. Methods for feature selection usually fall into two categories: wrappers and 'filters' (or a combination of the two) [Peng 2005, Liu 2005, Guyon 2003]. The wrapper models evaluate the features through the performance of a specific classifier. In the 'filter'-based approaches, the features are selected (filtered) using measures that are independent of the classification algorithms, yielding comparable classification errors for different classifiers. Filter approaches are computationally more efficient and offer a better generalization since they are independent of machine learning algorithms. Since we want to minimize the calibration time, the choice fell on filter methods. Two methods, r-square, r^2, and minimum redundancy maximum relevance (mRMR) were used in most of the experiments. Both methods are based on sequential feature selection (SFS).

4.3.1 r^2 correlation

The r^2 coefficient is a second order statistic measure computed from the square of the Pearson correlation coefficient r, that obtains the relevance of the features for a two-class discrimination. Considering samples of two classes X and Y, coefficient r is obtained

from

$$r(X,Y) = \frac{\sigma_{X,Y}}{\sigma_X.\sigma_Y} = \frac{\sum_{k=1}^{K}(X_k - \overline{X})(Y_k - \overline{Y})}{\sqrt{\sum_{k=1}^{K}(X_k - \overline{X})^2}\sqrt{\sum_{k=1}^{K}(Y_k - \overline{Y})^2}}. \quad (4.2)$$

The r^2 coefficient returns a value ranging from 0 to 1. High r^2 values denote large inter-class variance and small within-class variances. The r^2 is applied to each individual feature returning a score that allows to sequentially rank the features according to their discriminative power. Other measures such as Fisher score and Student's t-statistic were tested with similar results. Fig. 4.2 illustrates an example of the time instants of the 40 best ranked features selected with r^2 superimposed over the average of 90 P300-target epochs (using a dataset from LSC paradigm (chapter 6)). The same concept is extended to spatio-temporal features.

4.3.2 Minimum redundancy maximum Relevance

A different method called minimum redundancy and maximum relevance (mRMR) [Peng 2005] was also tested. This method tries to remove redundancy of the selected features. Two criteria are simultaneously applied: relevance and redundancy. Relevance gives a measure of correlation between the features and class labels while redundancy measures the statistical dependence between features. In the original method, both relevance

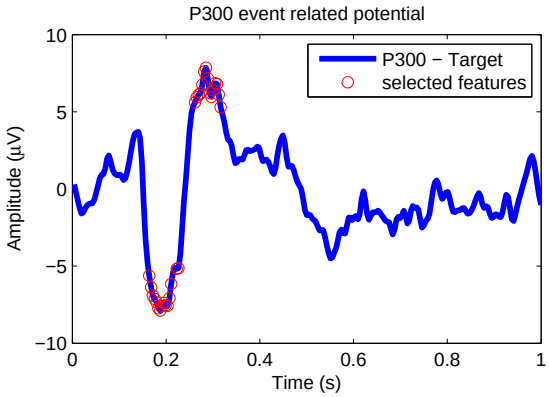


Figure 4.2: Average of 90 P300 target epochs (raw signal without any preprocessing) acquired at channel Pz. Circles represent the time instants of the 40 best ranked features selected with r^2 feature selector, superimposed over the amplitude of the time samples of the ERP.

and redundancy are computed through mutual information (MI). We slightly re-arranged mRMR, using MI for redundancy measurement and r^2 for feature relevance.

The optimum simultaneous maximization of relevance and minimization of redundancy is obtained by maximizing the criterion

$$\max C(S)/R(S) \tag{4.3}$$

where $R(S)$ are the MI values computed for each combination of feature subsets within a set of features S and $C(S)$ are the r^2 correlation values computed for each feature. In practice, mRMR method is implemented as a sequential suboptimal approach. It first ranks individual features according to relevance, from the highest to the lowest r^2 value, and then criterion (4.3) is achieved by iteratively selecting each feature according to

$$\max_{x_l \in \Omega - S_{m-1}} \left[\frac{C(x_l)}{\frac{1}{m-1} \sum_{x_k \in S_{m-1}} I(x_l, x_k)} \right]. \tag{4.4}$$

The initial set Ω of features is divided into two subsets: a selected features set (S_{m-1}) with $m-1$ features and the remaining feature set ($\Omega - S_{m-1}$). The first selected feature is the one with the largest relevance. Next, the average redundancy between the selected features and the remaining features is computed. The new selected feature is the one which maximizes (4.4) and it is iteratively selected from ($\Omega - S_{m-1}$). Although this iterative process is computationally efficient, it is naturally more demanding than the r^2 feature selector. We made a systematic analysis using several datasets obtained from some of the proposed paradigms, which showed that mRMR has a slightly higher performance than that of r^2, at the cost of an increased computation time [Pires 2009b]. However, for most of the online experiments, r^2 was the preferred method to reduce the time of the calibration phase.

4.4 Classification Methods

Many classification methods have been proposed in the context of P300-based BCIs. A generic BCI review is presented in [Lotte 2007] and a comparison of several P300 classification methods in [Krusienski 2006]. Some popular methods used in P300-based BCI are the linear discriminant analysis (LDA) or Fisher's linear discriminant (FLD)

[Lenhardt 2008, Hoffmann 2008b], the stepwise linear discriminant analysis (SWLDA) [Farwell 1988, Donchin 2000, Krusienski 2008, Townsend 2010], the support vector machines (SVM) [Rakotomamonjy 2008, Thulasidas 2006, Kaper 2004] and the neural networks (NN) [Cecotti 2011]. Although modern and more powerful classifiers, such as SVM and NN, show offline performance results sometimes above those achieved in state-of-the-art, they have very small representation in online experiments. This is mainly because model selection is very time consuming.

The detection of P300 evoked potentials is a binary classification problem, where P300 ERP associated to target events belongs to one class and EEG component associated with non-target events belongs to another class. Importantly, is that the binary classifiers provide scores for subsequent C-class symbol classifier (see section 3.1.2). Assuming $\mathbf{x}$ a feature vector, then a linear discriminant function is obtained from

$$g(\mathbf{x}) = \mathbf{w}.\mathbf{x} + b \tag{4.5}$$

where $\mathbf{w}$ is the weight vector and b is a bias term as depicted in Fig. 4.3 [Duda 2001]. To estimate the hyperplane that separates the two classes, a machine learning algorithm has to find $\mathbf{w}$ and b according to some optimization criterion. For a nonlinear separation, the above equation is modified according to

$$g(\mathbf{x}) = \mathbf{w}.\varphi(\mathbf{x}) + b \tag{4.6}$$

where $\varphi(\cdot)$ is a kernel that transforms the feature vector $\mathbf{x}$ into a new dimensional space. Classifiers can be broadly divided into generative and discriminative [Lotte 2007]. Generative classifiers learn the class models (e.g., LDA) while discriminative classifiers learn the way to discriminate different output values (e.g., SVM and NN). These classifiers are succinctly introduced below. Good general insights to machine learning, covering a wide range of pattern classification can be found in the books of Duda [Duda 2001] and Fukunaga [Fukunaga 1990].

4.4.1 Linear discriminant analysis

Under the restrictive assumptions of Gaussian distribution and equal covariance matrices of the two classes, LDA is shown to be Bayes optimal [Duda 2001], i.e. it minimizes the risk of misclassification of new samples drawn from the same distributions. Our own experience indicates that these assumptions are nearly satisfied for most of the ERPs

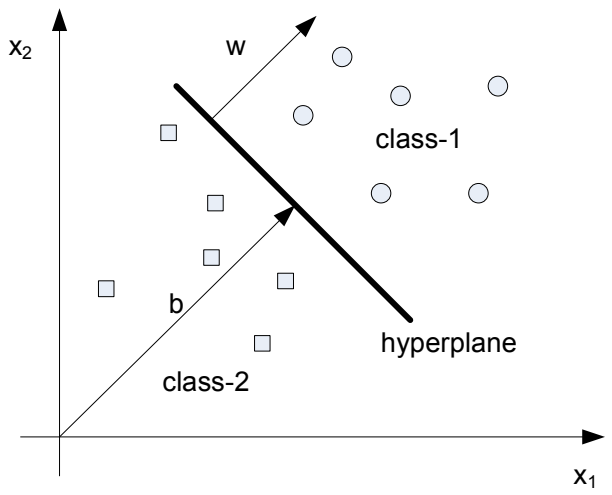


Figure 4.3: Linear decision hyperplane, where $\mathbf{y} = \mathbf{w}.\mathbf{x} + b = 0$ separates the feature space into class-1 ($\mathbf{y} > 0$) and class-2 ($\mathbf{y} < 0$).

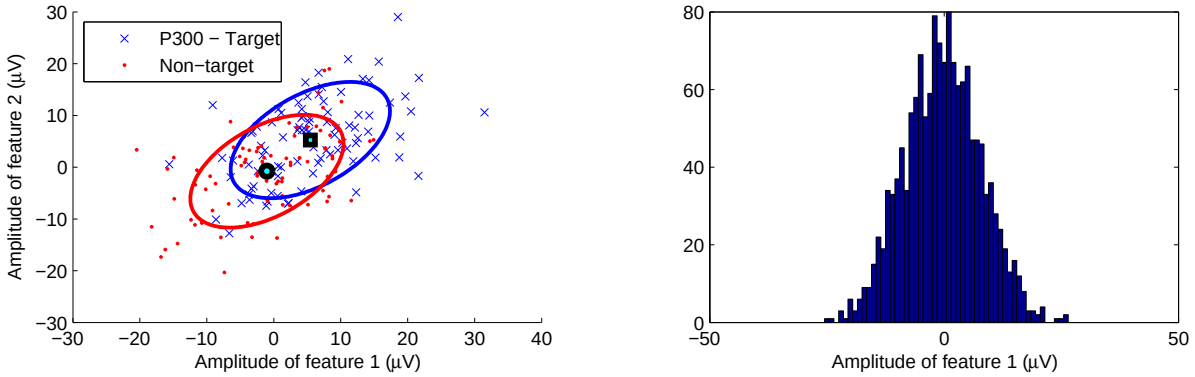
datasets as seen in Fig. 4.4, and as verified in other works [Blankertz 2011]. Under these assumptions it can be shown also that LDA is equivalent to FLD and to least square regression [Duda 2001]. Considering the feature vector distribution $p(\mathbf{x}) \sim \mathcal{N}(\mu, \Sigma)$, then the optimal decision boundary has the form $\mathbf{y} = \mathbf{w}.\mathbf{x} + b = 0$ where the weight vector is obtained from

$$\mathbf{w} = \Sigma^{-1}(\mu_1 - \mu_2) \tag{4.7}$$

where Σ is the covariance matrix and μ_1 and μ_2 are respectively the means of class-1 and class-2. For FLD, $\mathbf{w}$ is obtained by replacing Σ by $\Sigma_w = \Sigma_1 + \Sigma_2$, which represents the within-class scatter matrix [Duda 2001]. LDA and FLD are computationally very efficient.

4.4.1.1 Regularized LDA

When the number of training samples is small compared to feature dimensionality, overfitting problems may occur. If the number of features exceeds the number of samples, the covariance estimates do not have full rank and therefore can not be inverted. As we have seen in (4.1), the dimension of the feature vector can be very high. On the other hand, as one wants to minimize the BCI calibration time, the amount of gathered data is small. Additionally, the presence of muscular artifacts or strong noise induce the presence of outliers in the datasets. Regularization can prevent overfitting and the influence of outliers, and controls the complexity of the models. A regularized classifier provides higher generalization and robustness to outliers. One common approach of regularization is to



a) b)

Figure 4.4: a) Scatter plot of 2 best ranked features, for 90 target and non-target ERPs obtained from a dataset of LSC paradigm (chapter 6). Class covariances are indicated by the ellipse contour assuming Gaussian distributions (class-target with blue color and class-nontarget with red color) and class means marked by bold points. The classes have similar covariances; b) Histogram showing approximate Gaussian distribution of one best ranked feature, obtained from 900 non-target epochs of the same dataset used in a).

find a better estimate of Σ. We will use a shrinking approach derived from [Duda 2001] and arranged according to [Blankertz 2011]

$$\tilde{\Sigma} = (1-\gamma)\Sigma + \gamma\nu\mathbf{I} \tag{4.8}$$

where $\gamma \in [0,1]$ is a tuning parameter, and ν is given by $trace(\Sigma)/d$ of Σ with d being the feature dimensionality. This method is referred as regularized LDA (RLDA). When $\gamma = 0$, RLDA transforms into unregularized LDA.

4.4.2 Stepwise linear discriminant analysis

SWLDA is one of the most popular methods used in P300 systems and with effective online performance results. SWLDA is a LDA classifier that simultaneously selects the features to be included in function (4.5) [Krusienski 2008]. The weight vector $\mathbf{w} = [w_1 \quad w_2 \quad \cdots \quad w_{NT}]$ is obtained from multilinear least-square regression. A forward and backward stepwise regression is used to incrementally add or remove terms of the model, according to their statistical significance. A term w_i is included in the model if the p-value (of an F-statistic) of the model is below a given threshold (th_{in}). After a term is included into the discriminant function, a backward regression is performed to remove the least significant terms, having a p-value above a certain threshold (th_{out}). Each term

of $\mathbf{w}$ is set to zero if it is removed. The process is repeated until a predefined number of terms is achieved or when the entry/removal of terms does not improve the model. The two thresholds are the conditions to entry or remove the features. The features associated with non-zero terms are the selected features.

4.4.3 Naïve Bayes

The naïve Bayes (NB) classifier is a particular case of the rule-based Bayes classifier, which assumes independence between features. Despite this assumption, not actually satisfied in most cases, NB model is easy to fit and works surprisingly well outperforming other powerful methods in many situations [Rish 2001, Peng 2005, Ding 2005]. In a Bayes classifier, the attribution of a vector $\mathbf{x}$ to a class w_i, $i \in \{1, 2\}$ is made according to Bayes rule

$$P(w_i|\mathbf{x}) = \frac{P(w_i)p(\mathbf{x}|w_i)}{p(\mathbf{x})} \tag{4.9}$$

with

$$p(\mathbf{x}) = \sum_i p(\mathbf{x}|w_i)P(w_i) \tag{4.10}$$

where $P(w_i|\mathbf{x})$ is called the posterior conditional probability of class w_i, $p(\mathbf{x}|w_i)$ is the conditional probability, and $P(w_i)$ is the prior probability. The learning problem consists in estimating the $p(\mathbf{x}|w_i)$ from the training samples. The conditional density function of class w_i is modeled as a multivariate distribution under Gaussian assumption

$$p(\mathbf{x}|\mu_i, \Sigma_i) = \frac{1}{(2\pi)^{d/2}|\Sigma_i|^{1/2}} \exp(-(\mathbf{x} - \mu_i)^T \Sigma_i^{-1}(\mathbf{x} - \mu_i)/2) \tag{4.11}$$

where μ_i and Σ_i are the mean and covariance matrices computed for each class w_i from the training dataset. The posterior probability is usually computed taking the logarithmic transformation

$$\log(\pi_i) - \frac{1}{2}\log|\Sigma_i| - \frac{1}{2}(\mathbf{x} - \mu_i)^T \Sigma_i^{-1}(\mathbf{x} - \mu_i) \tag{4.12}$$

where π_i is the prior probability $P(w_i)$. The discriminant function can be defined as

$$g(x) = P(w_1|\mathbf{x}) - P(w_2|\mathbf{x}). \tag{4.13}$$

For a linear classifier the covariances are assumed to be the same for all classes $\Sigma = \Sigma_i$. For different covariance matrix Σ_i for each class, the classifier assumes a quadratic form. If

Σ is diagonal, then the features are assumed independent and we have the NB assumption

$$p(\mathbf{x}|w_i) = \prod_{j=1}^{d} p(x_j|w_i) \tag{4.14}$$

where x_j are the independent features of $\mathbf{x}$. Assuming different Gaussian distributions, then NB assumes a quadratic form, hereafter denoted as qNB. Compared to full covariance, NB benefits from a more accurate estimation of the probability density function (mean and variance) using a small amount of training data, since it is estimated for individual features. Moreover, the method is computationally efficient and does not require regularization. The requirement of a small amount of training data, and a fast model selection are two important characteristics to reduce the calibration time. Despite these favorable characteristics, only few relevant works have used NB in BCI. e.g., [Kohlmorgen 2004] for MRP classification. The NB assumption can be extended to the spatial domain by fusing the information of several channels as we successfully made in [Pires 2008b], generically expressed as

$$p(\mathbf{X}|w_i) = p(\mathbf{x}^{(1)} \wedge \mathbf{x}^{(2)} \wedge \cdots \mathbf{x}^{(N)}|w_i) = \prod_{j=1}^{N} p(\mathbf{x}^{(j)}|w_i) \tag{4.15}$$

where $\mathbf{x}^{(j)}$ is the feature vector associated to channel $j \in \{1 \cdots N\}$.

4.4.4 Support vector machines

While the above methods search to maximize the margin between the class means (generative classifiers), support vector machines (SVM) maximizes the margin between features (discriminative classifier) [Vapnik 1995, Burges 1998]. Assuming training samples $\mathbf{x}_k$ with class labels $y_k \in \{+1, -1\}$, then equation (4.6) can be rewritten as

$$y_k(\mathbf{w}.\varphi(\mathbf{x}_k) + b) \geq 1 - \xi_k \tag{4.16}$$

where $\xi_k \geq$ is a slack variable to cope with overlapping classes (soft constraint). The margin is given by $2/\|\mathbf{w}\|$, therefore to maximize the margin, the SVM optimization

problem is given by

$$\min_{\mathbf{w},b} \quad \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_k \xi_k \tag{4.17}$$

$$\text{s.t.} \quad y_k(\mathbf{w}.\varphi(\mathbf{x}_k) + b) \geq 1 - \xi_k \quad \text{for} \quad k \in \{1 \cdots K\} \tag{4.18}$$

where C is a tradeoff parameter (regularization parameter) that controls the compromise between having a large margin with more misclassifications, or having a small margin with less misclassifications, and K is the number of training samples. A strength of SVM is that it can be transformed into a nonlinear classifier in an effective way. To construct nonlinear boundaries, Kernels can be used transforming input features to a new features space [Boser 1992]. During our experiments, we used only the Gaussian Kernel (radial basis function, RBF)

$$K(\mathbf{x}_k, \mathbf{x}_j) = e^{\frac{-\|\mathbf{x}_k - \mathbf{x}_j\|^2}{2\sigma^2}}. \tag{4.19}$$

This results in a nonlinear discriminant function

$$g(\mathbf{x}) = \sum_k^{K_s} y_k \alpha_k K(\mathbf{x}, \mathbf{x}_k) + b \tag{4.20}$$

where K_s is the number of support vectors, and α_k are the Lagrangian multipliers used to solve the SVM optimization problem. The parameters C and σ were selected through grid search and checked with cross-validation. The learning algorithm to find the optimal hyperplane was the sequential minimal optimization (SMO) [Platt 1999]. Solving these problems can be very time consuming, especially when a large number of training examples is used.

4.5 Conclusion

This chapter presented several relevant issues that have to be addressed in BCI classification. Pre-processing and feature extraction are essential to increase SNR, and therefore they play an important role to the performance and computational optimization of the subsequent classifier. Feature extraction will be extensively addressed in chapter 5. Machine learning classifiers are undoubtedly a crucial step. The underlying concepts of some techniques usually used in BCI and in particular in P300-based BCIs, were here described. Our own experience shows that linear and simple algorithms are sufficient for P300 classification, as long as appropriate feature extraction methods are applied. This opinion

is supported by many other researchers [Krusienski 2006, Blankertz 2011], and explains why some methods such as SVM are sparsely used in practice in online BCIs. A comparative analysis made in the following chapter, section 5.5.1, clearly shows that with optimum spatial filtering, simple linear classifiers provide performances similar to SVM, without the increased complexity and computational cost of parameter tuning. Linear classifiers are less flexible, but more robust than nonlinear ones, since they have fewer free parameters to tune. For subject dependent models, the time to tune these parameters is time consuming leading to long calibration periods. Moreover, generative classifier need less training data [Martens 2010] and algorithms are easier and more effective to adapt to dynamic changes of neurophysiologic features of subjects and therefore they can be more easily extended to unsupervised learning [Daly 2008].

Chapter 5

Statistical Spatial Filtering

Contents

"First you guess. Don't laugh, this is the most important step. Then you compute the consequences. Compare the consequences to experience. If it disagrees with experience, the guess is wrong. In that simple statement is the key to science. It doesn't matter how beautiful your guess is or how smart you are or what your name is. If it disagrees with experience, it's wrong. That's all there is to it."

Richard Feynman (Nobel prize in Physics in 1965), Lecture to his students

In this chapter we propose and describe several novel statistical spatial filtering methods, and we show their importance for classification of event related potentials. Currently, most of the P300 based systems use a so-called classical spatio-temporal classification. Features are extracted from the time series of each channel and then concatenated forming an uni-dimensional feature vector used for classification (see equation (4.1)). In most cases, feature extraction is simply a process of decimation to reduce the dimension of the feature vector. We use here an alternative approach, in which EEG channels are spatially filtered in an optimum statistical sense, and then the spatial projections form the feature vector for classification. From a pattern recognition perspective, a spatial filter is itself a feature extractor applied in the spatial domain. The spatio-temporal and spatial approaches are schematically illustrated in Fig. 5.1. A comparison between them is given in section 5.5. The advantages and strengths of spatial filtering approach are supported both from a classification point of view, as well as from a neurophysiological perspective. Results achieved in an extensive set of experimental sessions (referred as *Case-study I*), and in tests with benchmark datasets, are compelling and fully validate the proposed methods.

5.1 Spatial filtering

5.1.1 State-of-the-art

There are three spatial filtering methods commonly applied in BCI: independent component analysis (ICA), principal component analysis (PCA) and common spatial patterns (CSP). Both ICA and PCA are mainly used on an unsupervised way: the former for separation of multichannel EEG data into statistically independent components, and the second for dimensionality reduction [Lenhardt 2008] and denoising. Most of the ICA applications have been in offline neurophysiologic analysis [Makeig 1999], and for strong artifact removal, such as eye blinking, eye movement and muscular activity [Jung 2000, Müller 2004]. Still, there are successful online and offline applica-

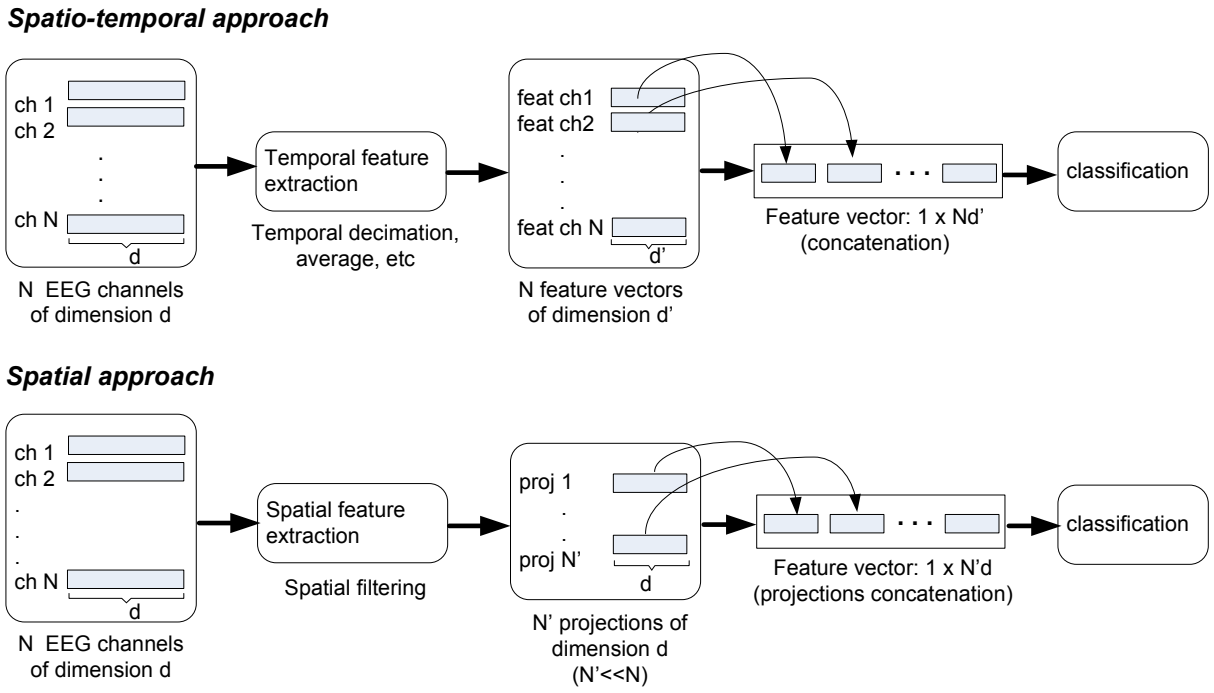


Figure 5.1: Spatio-temporal approach vs. spatial approach. In the spatio-temporal approach the reduction of the dimensionality of feature space is made in the temporal domain while in the spatial approach the reduction of the dimensionality of the feature space is made mainly in the spatial domain (see text in section 5.5 for more detailed information).

tions of ICA in the context of P300-based systems, as one can see respectively in [Serby 2005, Piccione 2006] and [Xu 2004]. The CSP method is a supervised technique that relies on the simultaneous diagonalization of two covariance matrices, maximizing the differences between two classes [Fukunaga 1970]. It was first applied in [Soong 1995] for localization of neurophysiologic features and since then it has been mainly applied in motor imagery based BCIs [Müller-Gerking 1999, Ramoser 2000, Blanchard 2004, Li 2004, Lemm 2005, Dornhege 2006], outperforming ICA and classical EEG re-referencing montages such as Laplacian derivations [Naeem 2006]. As concerns the effective use of CSP for P300 detection, see [Krusienski 2007] for a variant of CSP called common spatio-temporal patterns (CSTP) and [Pires 2009a] where a straightforward application of CSP was proposed[1]. In [Rivet 2009] it is proposed the xDAWN algorithm, which estimates spatial filters that find the evoked subspace by maximizing the signal-to-signal plus noise ratio. In other contexts than BCI, other spatial filtering techniques have been proposed specifically for ERP denoising [de Cheveigne 2008, Ivannikov 2009].

[1]CSP was our initial approach that gave rise to the proposed spatial filters presented in section 5.1.2. Therefore the implementation of CSP is given in Appendix B.1

5.1.2 Proposed methodologies

The proposed spatial filters are developed in a statistical beamformer framework based on eigenvectors. Beamforming techniques originally appeared in the field of antenna and sonar array signal processing [Van Veen 1988, Trees 2002] and are currently used in many other areas including magnetoencephalography (MEG) and EEG source reconstruction/localization [Van Veen 1997, Sekihara 2001, Grosse-Wentrup 2009].

Firstly, we propose a beamformer based on the classical SNR maximization criterion (Max-SNR) [Van Veen 1988]. The filter is obtained from the eigenvector that maximizes the output ratio of signal and noise powers. The method works blindly, i.e., it does not use geometrical information about the sensor array or the underlying sources. It requires the estimation of covariances matrices associated with periods of the P300 signal, and periods of only noise-plus-interference. Secondly, a beamformer based on the Fisher Criterion (FC) is proposed following the same eigenvector-based principle used in Max-SNR. The method extends the well known Fisher linear discriminant (FLD) to the spatial domain using an approach similar to [Hoffmann 2006]. Third, the two beamformers are cascaded in order to satisfy simultaneously in a suboptimum way both criteria [Fukunaga 1990, Ch.10]. This spatial filter is henceforth designated C-FMS (cascade Fisher and maximum signal-to-noise ratio). Forth, the FC beamformer is extended to the frequency domain leading to a method that combines simultaneously temporal and spatial filtering, which is designated by spatio-spectral Fisher criterion beamformer (SSFCB).

Max-SNR, FC and C-FMS spatial filters are experimentally assessed in this chapter in the context of an experimental study made with the standard RC speller paradigm (*Case-study I*), while SSFCB is experimentally assessed in the context of two other paradigms presented in chapters 6 and 8. Their offline performances are compared with those obtained with best channel and with Laplacian spatial filtering. Two $C-class$ methodologies are compared offline, one combining the average of the signal epochs and the other combining the *a posteriori* probabilities. Online validation of the entire system methodology was obtained by measuring the ITR, SPM and accuracy, in a case study with 19 able-bodied participants and 5 disabled participants. The online experiments used the C-FMS spatial filter, and the epoch average combination for the $C-class$ character recognition. The system requires a very short calibration time of about 7 minutes, specifically, 5 minutes to collect labeled data (calibration session) plus less than 2 minutes to obtain spatial filters and classification models (Bayesian classifier).

Further validation using benchmark data sets of the BCI-competition 2003 [BCI-Competition 2003] is provided for state-of-the-art comparison of the proposed

methodology.

5.2 Model: Assumptions and notation

An ERP is modeled as a random (stochastic) process from which we know a collection (ensemble) of realizations of that process (Fig. 5.2). The realizations are obtained by collecting the neuronal responses (epochs) to a given stimulus event. The process is considered wide-sense stationary, i.e., assuming that the first and second-order statistics (mean and variance) of the process are time-invariant. Moreover, it will be assumed that the process follows a Gaussian distribution, which is generically defined for a scalar-valued random variable x as

$$\mathcal{N}(x, \mu_x, \sigma_x) = \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left[-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right] \tag{5.1}$$

where μ_x and σ_x are the mean and the variance of the random variable x.

Let us start modeling our process, considering an EEG epoch $\mathbf{X}$ defined as a time sequence of measures, $\mathbf{X} = [\mathbf{x}(t_1) \quad \mathbf{x}(t_2) \quad \cdots \quad \mathbf{x}(t_T)]$, where T is the number of time samples and $\mathbf{x}(t)$ is a column vector with dimension N (number of EEG channels). Therefore, each epoch is represented by a spatio-temporal matrix $\mathbf{X} \in \mathbb{R}^{N \times T}$ with dimension $N \times T$ (in our case, $N = 12$ channels and $T = 256$ samples). Target and non-target epochs are represented respectively by $\mathbf{X}_+$ and $\mathbf{X}_-$, where the subscripts $_+$ and $_-$ stand respectively for target and non-target.

Let us consider the target epochs modeled according to

$$\mathbf{X}_{+,k} = \mathbf{S}_k + \mathbf{V}_k \tag{5.2}$$

where $\mathbf{X}_{+,k}$ is the k_{th} recorded epoch and $\mathbf{S}_k$ is the k_{th} P300 signal, measured in the N dimensional space. $\mathbf{V}_k$ contains activity from ongoing EEG, plus the interference from not-attended flashes, plus white noise. Non-target epochs occur immediately before target epochs and thus the activity should be similar to $\mathbf{V}_k$. Hence, $X_{-,k}$ is modeled as the noise and interference part of the measured target epochs

$$\mathbf{X}_{-,k} = \mathbf{V}_k. \tag{5.3}$$

Models (5.2) and (5.3) were experimentally sustained by means of a frequency analysis. It consisted of calculating and analyzing the FFT spectra over representative data

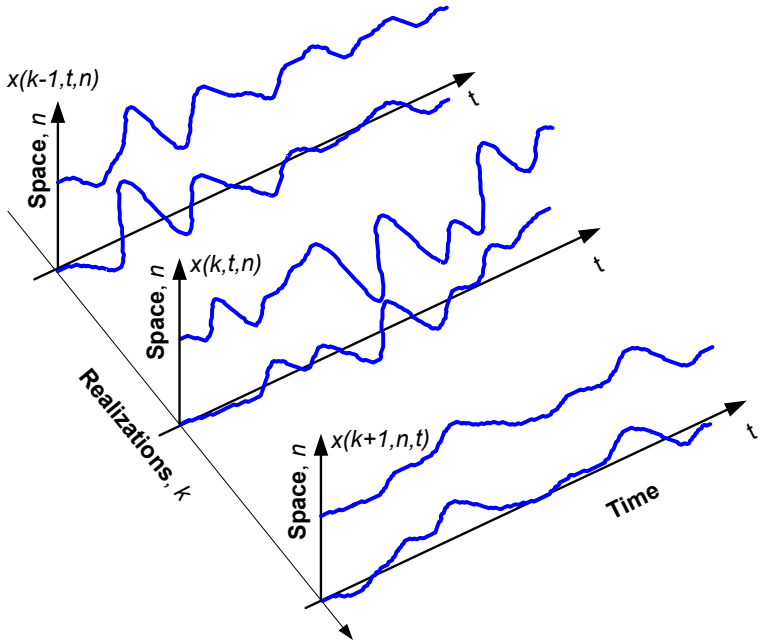


Figure 5.2: Ensemble of realizations (epochs) of a random process. A temporal and spatial view over a set of realizations. The symbols t, n and k stand respectively for time, channel and realization.

collected from one session (180 target epochs and 900 non-target epochs) with RC speller. Color maps in Fig. 5.3(a) and Fig. 5.3(b) represent respectively the FFT spectra of 90 target and 90 non-target epochs measured at channel Pz. The spectra for both conditions, $\mathbf{X}_+$ and $\mathbf{X}_-$, present similar frequency distributions. This overlapping of spectra is evidenced in the example of a single realization in Fig. 5.3(c). This shows, firstly, that much of the non-target activity is contained in target epochs, and secondly, that temporal filtering is insufficient to remove noise from target epochs, and thus it should be used carefully. Figure 5.3(d) presents the average of the FFT spectra of target and non-target epochs. The average attenuates the spectrum of random components, and emphasizes the spectrum of the P300 ERP and other uncorrelated interfering signals. A strong interference at 5 Hz appears in the target spectrum (see Fig. 5.3(d)). This interference comes from the rows/columns of the RC speller flashing with an SOA of 200 ms, i.e, 5 Hz (see its effect in time domain in Fig. 5.4). These stimuli generate a steady state visual evoked potential (SSVEP) at 5 Hz, and a 2nd harmonic at 10 Hz as well. This 2nd harmonic affects target epochs with less impact because it does not overlap the spectrum of the P300 ERP.

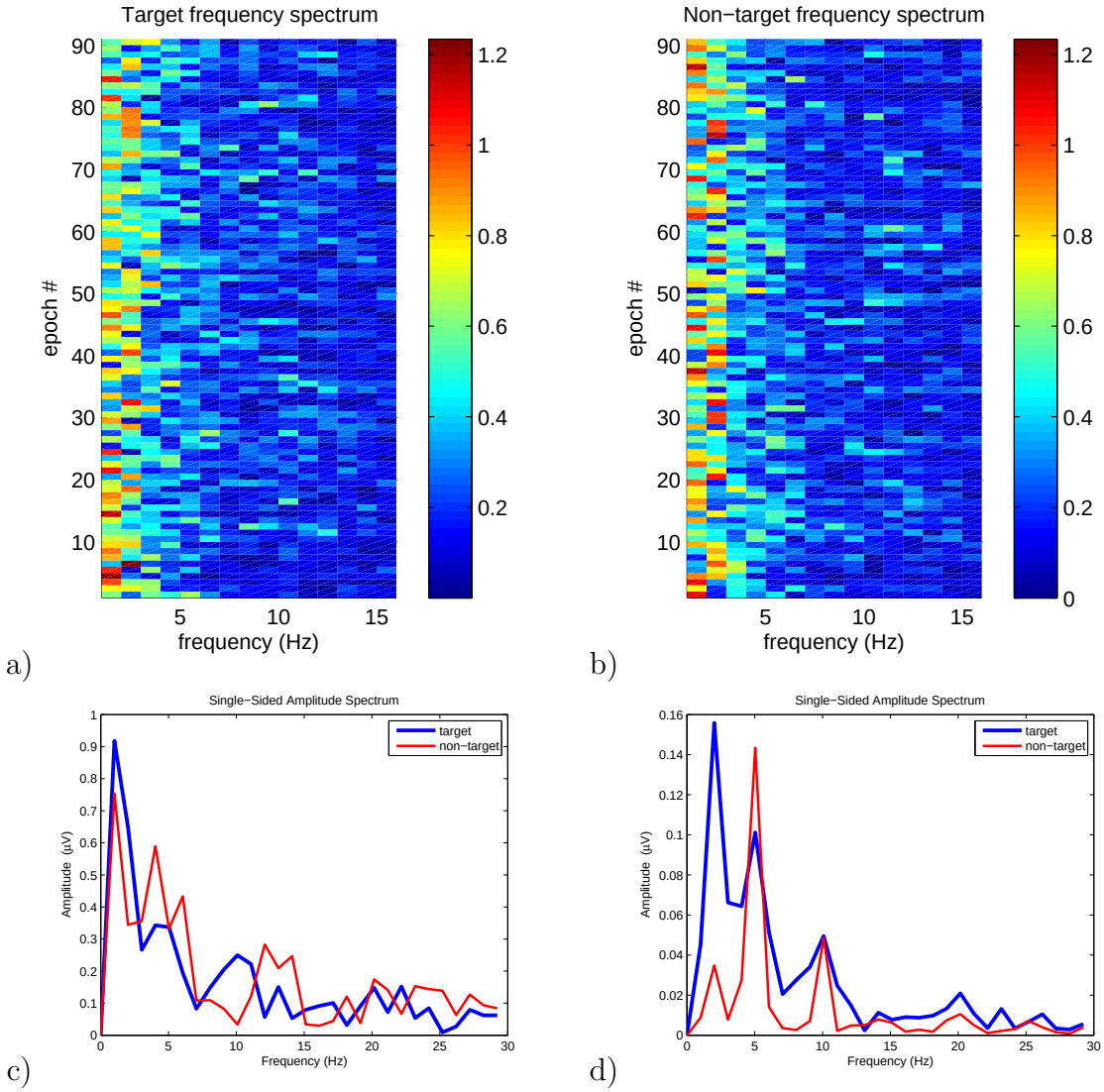


Figure 5.3: FFT spectrum of a representative data set of one session (180 target and 900 non-target epochs) measured at channel Pz; (a) Color map of the FFT spectra over 90 out of the 180 target epochs; (b) Color map of the FFT spectra over 90 out of the 900 non-target epochs; (c) Example of one FFT of a single epoch (target and non-target); (d) Average of the FFT spectra of all epochs (180 FFTs of target epochs and 900 FFTs of non-target epochs).

5.3 Statistical spatial filters

A spatial filter is generically an weighting vector, w, that combines the data of N channels at each time instant t

$$\mathrm{y}_j(t) = \sum_{i=1}^{N} \mathrm{w}_{ij}\mathrm{x}_i(t) \quad , j = 1, \cdots, N \tag{5.4}$$

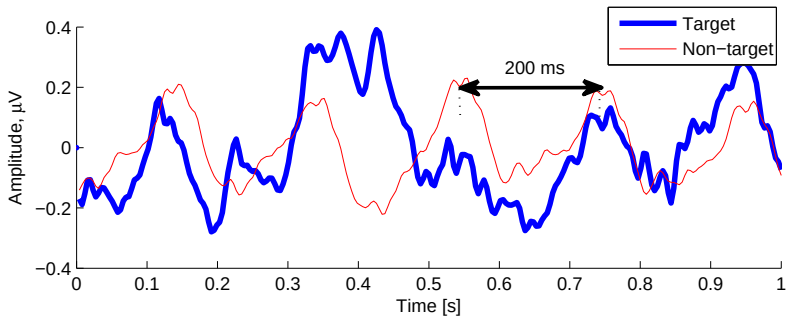


Figure 5.4: Average of 180 unfiltered target epochs and 900 unfiltered non-target epochs recorded at channel Pz. The result evidences an oscillatory component with 200 ms period in non-target epochs, which is also visible in target epochs.

where y_j is the output projection obtained from input channels x_i, which can be denoted in the matrix notation from

$$\mathbf{Y} = W'\mathbf{X} \tag{5.5}$$

where $'$ denotes the transpose operator, or in the expanded form

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix} = \begin{bmatrix} W_{11} & W_{12} & \cdots & W_{1N} \\ W_{21} & W_{22} & \cdots & W_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ W_{N1} & W_{N2} & \cdots & W_{NN} \end{bmatrix}' \cdot \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix}. \tag{5.6}$$

5.3.1 Max-SNR beamformer

The application of eigenvectors in statistical signal processing for low-rank modeling, i.e., for reducing the feature data space while keeping the intrinsic information, has been used to determine optimum temporal finite impulse response (FIR) filters [Haykin 1996]. We apply, in this first approach, the same concept to the spatial domain, by stating the spatial filtering of P300 as a spatial denoising problem. The solution is an optimum beamformer, based on statistical data, that maximizes the output SNR

$$\mathrm{SNR} = \frac{\mathrm{E}[W'\mathbf{S}\mathbf{S}'W]}{\mathrm{E}[W'\mathbf{X}_{-}\mathbf{X}'_{-}W]} \backsimeq \frac{W'\overline{\mathbf{R}}_{+}W}{W'\overline{\mathbf{R}}_{-}W} \tag{5.7}$$

where W is the weighting vector, $\mathrm{E}[\cdot]$ represents the expectation operator, and the matrices $\overline{\mathbf{R}}_{+}$ and $\overline{\mathbf{R}}_{-}$ are the estimated covariance matrices for target and non-target. The

maximum SNR is obtained by maximizing the discriminative Rayleigh quotient in (5.7). The optimal W is the eigenvector associated to the largest eigenvalue. The solution is achieved by finding the generalized eigenvalue decomposition that satisfies the equation (see a formal demonstration in Appendix B.2)

$$\overline{\mathbf{R}}_{+}W = \overline{\mathbf{R}}_{-}W\Lambda \tag{5.8}$$

where Λ is the eigenvalue matrix. The eigenvectors W are obtained from the eigenvalue decomposition of $(\overline{\mathbf{R}}_{-})^{-1}\overline{\mathbf{R}}_{+}$ provided that $\overline{\mathbf{R}}_{-}$ is nonsingular. The principal eigenvector $W^{(1)}$ maximizes the SNR, and therefore the output of the beamformer is given by

$$\mathbf{y} = W^{(1)\prime}\mathbf{X}. \tag{5.9}$$

The $N \times T-$dimensional measurement $\mathbf{X}$ is transformed into a 1-dimensional subspace, $1 \times T$. This reduction of the feature space is an important achievement for subsequent classification.

The matrices $\overline{\mathbf{R}}_{+}$ and $\overline{\mathbf{R}}_{-}$ are estimated from the average over the epochs within each class, gathered during calibration sessions. Consider the $N \times N$ normalized spatial covariance for each epoch k given by $\mathbf{R}_k = \mathbf{X}_k\mathbf{X}_k'/tr(\mathbf{X}_k\mathbf{X}_k')$, then, $\overline{\mathbf{R}}_{+}$ and $\overline{\mathbf{R}}_{-}$ are computed from

$$\overline{\mathbf{R}}_{+} = \langle \mathbf{R}_{+} \rangle = \frac{1}{K_{+}} \sum_{k=1}^{K_{+}} \mathbf{R}_{+,k} \quad \text{and} \quad \overline{\mathbf{R}}_{-} = \langle \mathbf{R}_{-} \rangle = \frac{1}{K_{-}} \sum_{k=1}^{K_{-}} \mathbf{R}_{-,k} \tag{5.10}$$

where K_{+} and K_{-} are the number of target and non-target training samples. Regularization of the covariance $\overline{\mathbf{R}}_{-}$ can be included according to

$$\overline{\mathbf{R}}_{+}W = [(1-\alpha)\overline{\mathbf{R}}_{-} + \alpha I]W\Lambda \tag{5.11}$$

where $\alpha \leq 1$, can alleviate overfitting and improve class discrimination.

The Max-SNR solution (5.8) is similar to that obtained from CSP, which can also be stated as a generalized eigenvalue problem [Tomioka 2007]. The Max-SNR beamformer can be regarded as a particular case of CSP (see Appendix B.3 for a formal demonstration).

5.3.2 FC Beamformer

The Max-SNR criterion relies on the ratio of signal and noise cross-powers. From a pattern recognition perspective, other criteria can be investigated to implement a beamformer. One of such criteria is the Fisher's criterion (FC) [Duda 2001], which aims to increase the separation between classes while minimizing the variance within a class (Fisher linear discriminant, FLD). This concept can be extrapolated to the spatial domain using spatio-temporal data as it was done in Max-SNR (section 5.3.1). The FC takes into consideration the difference between target and non-target spatio-temporal patterns. Then, it is expected that the spatial filter maximizes the spatio-temporal differences, leading to an enhancement of specific subcomponents of the ERP. The FC is given by the Rayleigh quotient

$$J(W) = \frac{W'\mathbf{S}_b W}{W'\mathbf{S}_w W} \tag{5.12}$$

where $\mathbf{S}_b$ is the spatial between-class matrix and $\mathbf{S}_w$ is the spatial within-class matrix. The optimum filter W is found solving the generalized eigenvalue problem

$$\mathbf{S}_b W = \mathbf{S}_w W \Lambda. \tag{5.13}$$

The selected filter is the eigenvector associated with the largest eigenvalue, i.e., $W^{(1)}$, and the spatial filter output is obtained by applying expression (5.9).

Taking the spatio-temporal matrix $\mathbf{X}_k$ (dimension $N \times T$) from each epoch k, the matrices $\mathbf{S}_b$ and $\mathbf{S}_w$ are computed from

$$\mathbf{S}_b = \sum_i p_i(\overline{\mathbf{X}}_i - \overline{\mathbf{X}})(\overline{\mathbf{X}}_i - \overline{\mathbf{X}})' \tag{5.14}$$

and

$$\mathbf{S}_w = \sum_i \sum_{k \in C_i} (\mathbf{X}_{i,k} - \overline{\mathbf{X}}_i)(\mathbf{X}_{i,k} - \overline{\mathbf{X}}_i)' \tag{5.15}$$

where $i \in \{+, -\}$ and, C_+ and C_- represent respectively the target and non-target classes, and p_i is the class probability. The average of the epochs in class C_i and the average of all epochs are respectively denoted $\overline{\mathbf{X}}_i$ and $\overline{\mathbf{X}}$, with

$$\overline{\mathbf{X}}_i = \frac{1}{K_i} \sum_{k=1}^{K_i} \mathbf{X}_{i,k} \quad \text{and} \quad \overline{\mathbf{X}} = \frac{1}{K} \sum_{k=1}^{K} \mathbf{X}_k \tag{5.16}$$

where K_i is the number of epochs in class C_i and K is the total number of epochs. To

increase generalization, $\mathbf{S}_w$ in (5.13) can be regularized according to

$$\mathbf{S}_b W = [(I-\theta)\mathbf{S}_w + \theta I]W\Lambda \tag{5.17}$$

where θ is the regularization parameter that can be adjusted from training data to increase class discrimination.

5.3.3 C-FMS beamformer

In order to satisfy both Max-SNR and FC, a cascade of the two spatial filters is proposed using a suboptimum approach [Fukunaga 1990, Ch.10]. FC is applied first since it is more discriminative than Max-SNR. The first transform is obtained from

$$\mathbf{Y} = W_1{}'\mathbf{X} \tag{5.18}$$

where W_1 is the spatial filter computed according to (5.17) (subscript $_1$ denotes the filter of the first stage of the cascade). The first feature vector is obtained from the first projection

$$\mathbf{y_1} = W_1^{(1)\prime}\mathbf{X}. \tag{5.19}$$

The feature vector $\mathbf{y_1}$ preserves the information about FC while the remaining components in the $(N-1)-$dimensional space are used for a second transform

$$\mathbf{Z} = W_2{}'\mathbf{Y}^{(2:N)} \tag{5.20}$$

where the spatial filter W_2 is estimated according to (5.11) taking $\mathbf{Y}^{(2:N)}$ data from the 1st stage of cascade (5.18). The first projection satisfies the Max-SNR criterion

$$\mathbf{z_1} = W_2^{(1)\prime}\mathbf{Z}. \tag{5.21}$$

The concatenation of the two projections

$$\mathbf{v} = [\mathbf{y_1} \quad \mathbf{z_1}] \tag{5.22}$$

maximizes both FC and max-SNR criteria in a suboptimum way.

5.3.4 Frequency domain SSFCB

The FC beamformer is here extended to the frequency domain combining simultaneously temporal and spatial filtering, leading to a new method henceforth called spatio-spectral Fisher criterion beamformer (SSFCB). The importance of frequency in P300-BCIs has been disregarded by most of the research community, yet, as seen in section 5.2, frequency is quite informative. The main goal of SSFCB is to embed a spectral frequency shaping into the spatial filter, which can be efficiently implemented if the overall method is formulated in the frequency domain. Spectral shaping introduces an increased flexibility to the feature extractor. It can be used simply as selective filter, working as a low-pass, high-pass or band-pass filter, or it can be used as a shaping filter, by selecting or weighting discriminative frequency features. The generic derivation of SSFCB follows the same eigenvector principle of the above spatial filters. The block diagram representing the overall implementation is shown in Fig. 5.5. Let us first consider signals in the frequency domain. For a finite discrete-time signal $x[n]$, of length T time samples, the discrete Fourier Transform (DFT) is defined as [Oppenheim 1999]

$$X(k) = \sum_{n=0}^{T-1} x[n] e^{-j(2\pi/T)nk} \quad k = 0, \cdots, T-1, \tag{5.23}$$

and the inverse discrete Fourier Transform (IDFT) is defined by

$$x[n] = \frac{1}{T} \sum_{k=0}^{T-1} X(k) e^{j(2\pi/T)nk} \quad n = 0, \cdots, T-1. \tag{5.24}$$

The index k corresponds to kf_s/T Hz, at a sampling rate f_s. In a matrix notation, the DFT equation can be computed as a linear transformation, $\mathbf{X} = F\mathbf{x}$, where the (k, n) elements of the transformation matrix $F \in \mathbb{R}^{T \times T}$ are $\frac{1}{\sqrt{T}} e^{-j(2\pi/T)nk}$. For a real input sequence, the DFT is conjugate symmetric, which means that the DFT elements are half redundant. Thus, the DFT can be expressed only by the first $T' = T/2$ frequency bins. The DFT is computed using the fast fourier transform (FFT), an efficient computation algorithm. For sake of concise notation and to avoid confusion with time domain methods, vectorial and matrix representation in the frequency domain will be denoted with the tilde symbol, i.e DFT$(\mathbf{x}) = \tilde{\mathbf{x}}$. The frequency domain representation of the spatio-temporal matrix of a single EEG epoch $\mathbf{X} \in \mathbb{R}^{N \times T}$, defined in section 5.2, is given by $\tilde{\mathbf{X}} \in \mathbb{R}^{N \times T'}$. The spectral filtering is performed in the frequency domain by applying the

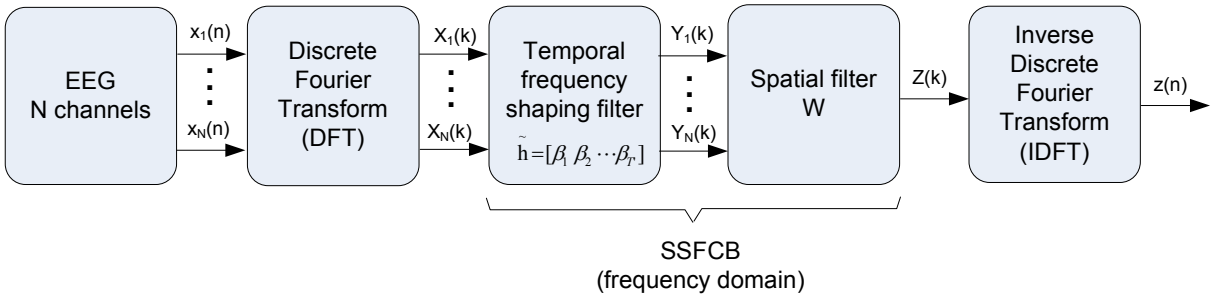


Figure 5.5: Block diagram of the implementation of SSFCB.

DFT convolution theorem to each channel i

$$\mathbf{x}^i \circledast \mathbf{h} = \tilde{\mathbf{x}}^i \cdot \tilde{\mathbf{h}} \tag{5.25}$$

where $\mathbf{h}$ represents the filter, '$\circledast$' denotes the convolution operation and '$\cdot$' denotes a pointwise multiplication. In the N dimensional domain, (5.5) is now stated as a spatio-spectral projection in the frequency domain

$$\tilde{\mathbf{Y}} = \tilde{W}^H \tilde{\mathbf{X}}.\tilde{\mathbf{h}} = \tilde{W}^H \tilde{\mathbf{Z}} \tag{5.26}$$

where the superscript H denotes the Hermitian transpose, and $\tilde{\mathbf{Z}} \in \mathbb{R}^{N \times T'}$ are the spectrally shaped components in the frequency domain. The spectral filter is an weighting vector $\tilde{\mathbf{h}} = [\beta(1) \quad \beta(2) \quad \cdots \quad \beta(T')]$. The coefficients β are adjusted according to the discriminative frequency bands. In its simple form, the weights act as a mask where 1 is used for a discriminative frequency and 0 for a non-discriminative frequency. The spatial filter $\tilde{W}$ is obtained by maximizing the Rayleigh quotient

$$J(\tilde{W}) = \frac{\tilde{W}^H \tilde{\mathbf{S}}_b \tilde{W}}{\tilde{W}^H \tilde{\mathbf{S}}_w \tilde{W}}. \tag{5.27}$$

Matrices $\tilde{\mathbf{S}}_b$ and $\tilde{\mathbf{S}}_w$ are obtained following the expressions (5.14), (5.15), (5.16), and replacing $\mathbf{X}$ by $\tilde{\mathbf{Z}}$ and the transpose operator $'$ by the complex Hermitian transpose H. The first projection of the spatio-spectral filter is obtained in the frequency domain from

$$\tilde{\mathbf{y}} = \tilde{W}^{(1)\prime} \tilde{\mathbf{Z}} \tag{5.28}$$

and in the time domain from $\mathbf{y} = \text{IDFT}(\tilde{\mathbf{y}})$. If coefficients β are all set to 1, $\tilde{\mathbf{h}} = [1 \quad 1 \quad \cdots \quad 1]$, then there is no spectral shaping, and the method behaves as the spatial filter in the time domain. In this case, spatial projections in the frequency domain coincide

with those obtained in the time domain, since the DFT is a linear operator, and the scatter matrices $\tilde{\mathbf{S}}_b$ and $\tilde{\mathbf{S}}_w$ preserve the energy according to Parseval's relation

$$\sum_{n=0}^{T-1} |x(n)|^2 = \frac{1}{T} \sum_{k=0}^{T-1} |X(k)|^2 .$$

5.3.4.1 Spectral shaping coefficients

Coefficients of the spectral filter $\tilde{\mathbf{h}} = [\beta(1) \quad \beta(2) \quad \cdots \quad \beta(T')]$ can be selected manually by setting the values to ones or zeros, or can be tuned automatically. The manual selection is equivalent to a on/off filter bank, in which, pre-defined non-discriminative bands are set to 0. An algorithm for automatic selection of the coefficients, that adjusts simultaneously spectral and spatial filters (automatic SSFCB), is discussed in chapter 8.

5.4 *Case-study I* - Spatial filtering approach in the context of the RC speller BCI

The performance of Max-SNR, FC and C-FMS spatial filters were experimentally assessed offline, and C-FMS was further assessed online, all in the context of an experimental study made with the standard row-column speller paradigm. The experiments were performed by able-bodied and motor impaired persons. A summary of the experimental protocol of *case-study I* is in Table 5.1, and a succinct description is given in the following sections.

5.4.1 Participants

Nineteen able-bodied volunteers, three subjects with CP, and two subjects with ALS participated in this study. Fourteen of the able-bodied subjects and the five disabled subjects never had used a BCI before. Table 5.2 presents a summary of clinical data of disabled subjects, as well as a description of their assistive interfaces used in their daily lives. The three subjects with CP are confined to a wheelchair. Subject S20 steers the wheelchair using an head-switch that selects the direction via a scanning interface, subject S21 uses an adapted joystick controlled by the right foot, and subject S22 controls the wheelchair with the chin. Subjects S23 and S24 present a bulbar-onset ALS whose main signs are dysarthria and dysphagia. Subject S23 starts to evidence also muscular weakness in superior limbs with distal predominance. Spoken communication with subjects S20-S23 was hard, and impossible with subject S24. All patients presented normal cognitive

Table 5.1: *Case-study I*: summary of experimental protocol

Recording	
Channels:	12 Ag/Cl electrodes Fz, Cz, C3, C4, CPz, Pz, P3, P4, PO7, PO8, POz and Oz (10-20 extended positions) Reference: left or right ear lobe; Ground: AFz
Frequency sampling:	256 Hz
Preprocessing filters:	0.1-30 Hz bandpass filter and 50 Hz notch filter
Impedance:	$< 10K\Omega$
Paradigm protocol	
Paradigm:	6×6 matrix RC speller
RC parameters:	SOA=200 ms, ISI=100 ms, ITI=3.5 s
Screen:	15" at about 60-70 cm
Participants	19 able-bodied, 3 CP and 2 ALS

Table 5.2: *Case-study I* - summary of clinical data of motor impaired subjects, and interfaces currently used.

Subject/Sex	**Age**	**Diagnosis**	**Time since diagnosis**	**Main signs / Main functionality**	**Adapted interfaces and assistive devices**
S20/F	18	CP	Posnatal	Tetraparesis, dystonia with spasticity and dysarthria / Head control	Powered wheelchair: head-switch connected to a scanning interface; Computer: head-switch and head-tracking
S21/M	34	CP	Perinatal	Tetraparesis, dystonia with spasticity, and high dysarthria / Head and right foot control	Powered wheelchair: adapted joystick controlled by right foot; Computer: controls mouse and keyboard with the right foot
S22/M	46	CP and discal hernia C3-C4	Perinatal	Tetraparesis, spasticity and dysarthria / Head control	Powered wheelchair: joystick adapted to chin; Computer: head-tracking
S23/F	67	Bulbar-onset ALS (FRS-r 46)	7 years	Dysarthria, dysphagia, and muscular weakness in upper limbs / -	-
S24/F	75	Bulbar-onset ALS (FRS-r 40)	1 year	High dysarthria and dysphagia / -	-

capabilities. The group of able-bodied volunteers was composed of 10 males and 9 females with ages from 18 to 42 years old, averaging 30.1 years old.

5.4.2 Experimental protocol

Participants were asked to control the RC speller introduced in section 3.1. The experiments took place on regular rooms in an environment with some noise and people moving around, at APCC (CP subjects), HUC (ALS subjects) and at our laboratory facilities (able-bodied subjects). The experiments consisted of a calibration phase and an online phase. Before the calibration phase, the subjects were instructed to be relaxed and attend the desired target, mentally counting the number of intensifications of target rows and columns. The able-bodied and ALS subjects were seated on a standard chair, while the CP subjects were seated at their own wheelchairs. A 15" computer screen was positioned in front of the participants at about 60-70 cm. It was asked only to the able-bodied subjects to avoid blinking and moving the eyes. During the calibration phase, the subjects attended the letters of the word 'INTERFACE' (9 characters) which were successively provided at the top of the screen (Fig. 3.1). Each row and column was repeated 10 times for each letter. Therefore, the data collected during the calibration phase consisted of 180 target epochs and 900 non-target epochs (5 minutes calibration). The EEG activity was recorded with the g.tec gUSBamp amplifier according to setup in Table 5.1.

5.4.3 Offline classification results

For each participant, the classification models were obtained from one training data set collected during the calibration session. A second data set, with the same amount of data, was collected for testing, such that all offline results presented in this section were obtained from unseen data.

The classification performance was assessed using the NB classifier in its quadratic naïve form (qNB) (see section 4.4.3)

$$\begin{aligned} p(\mathbf{y}|C_i) &= \prod_{j=1}^{N_f} p(\mathbf{y}(j)|C_i) = \\ &= \prod_{j=1}^{N_f} \frac{1}{\sqrt{2\pi}\sigma_i(j)} \exp\left(-\frac{(y(j)-\mu_i(j))^2}{2\sigma_i^2(j)}\right) \end{aligned} \tag{5.29}$$

where each feature j is assumed to have a normal distribution $\mathcal{N}(\mu_i(j), \sigma_i^2(j))$. The number of features is defined by N_f, and C_i ($i \in \{+,-\}$) represents the target and non-target classes. The *a posteriori* probability $p(C_i|\mathbf{y})$ is computed from the conditional

probabilities using the Bayes theorem:

$$P(C_i|\mathbf{y}) = \frac{P(C_i)p(\mathbf{y}|C_i)}{p(\mathbf{y})}. \tag{5.30}$$

The prior probabilities $P(C_i)$ are respectively 2/12 and 10/12 for target and non-target. The class is detected using the following maximum *a posteriori* decision rule

$$\hat{c} = \arg\max\{P(C_+|\mathbf{y}), P(C_-|\mathbf{y})\}. \tag{5.31}$$

A trial is composed of several sub-trials, and thus the binary classifier must be applied combining the epochs of all sub-trials associated to each event (see section 3.1.2). Two approaches were followed. In the first, the spatial filter was applied to the average of the K-epochs associated to each event and then the *a posteriori* probability was obtained according to

$$P(C_i|\mathbf{y}) \equiv P(C_i|\frac{1}{K}\sum_{k=1}^{K}\mathbf{y_k}) \qquad , i \in \{+,-\}. \tag{5.32}$$

In the second approach, the spatial filter was applied to single epochs and then the K-posterior probabilities were combined according to

$$P(C_i|\mathbf{y}) \equiv \prod_{k=1}^{K} P(C_i|\mathbf{y_k}) \qquad , i \in \{+,-\} \tag{5.33}$$

where $P(C_i|\mathbf{y_k})$ is the *a posteriori* probability for the epoch k and K is the number of epochs (repetitions). Class detection was done in both cases using $P(C_i|\mathbf{y})$ in (5.31).

Figure 5.6a) shows the classification error rate following the K-epoch average approach. The error rate was obtained averaging the results of all 23 subjects (subject S21 did not elicit a traceable P300 and was discarded from the analysis), i.e., using $23 \times 180 = 4140$ target epochs and $23 \times 900 = 20700$ non-target epochs. The plot shows results of the 3 proposed spatial filters, and for sake of comparison, the results of Laplacian (LAP) derivations, as well as the results concerning the channel presenting the highest discrimination. The Laplacian method is a high-pass spatial filter that computes for each electrode the instantaneous second derivative of the spatial voltage distribution, emphasizing localized activity and attenuating surrounding activity [Srinivasan 1999]. It is an unsupervised technique that, in some cases, significantly increases the SNR and thereby increases the classification accuracy [McFarland 1997]. The Laplacian of channel

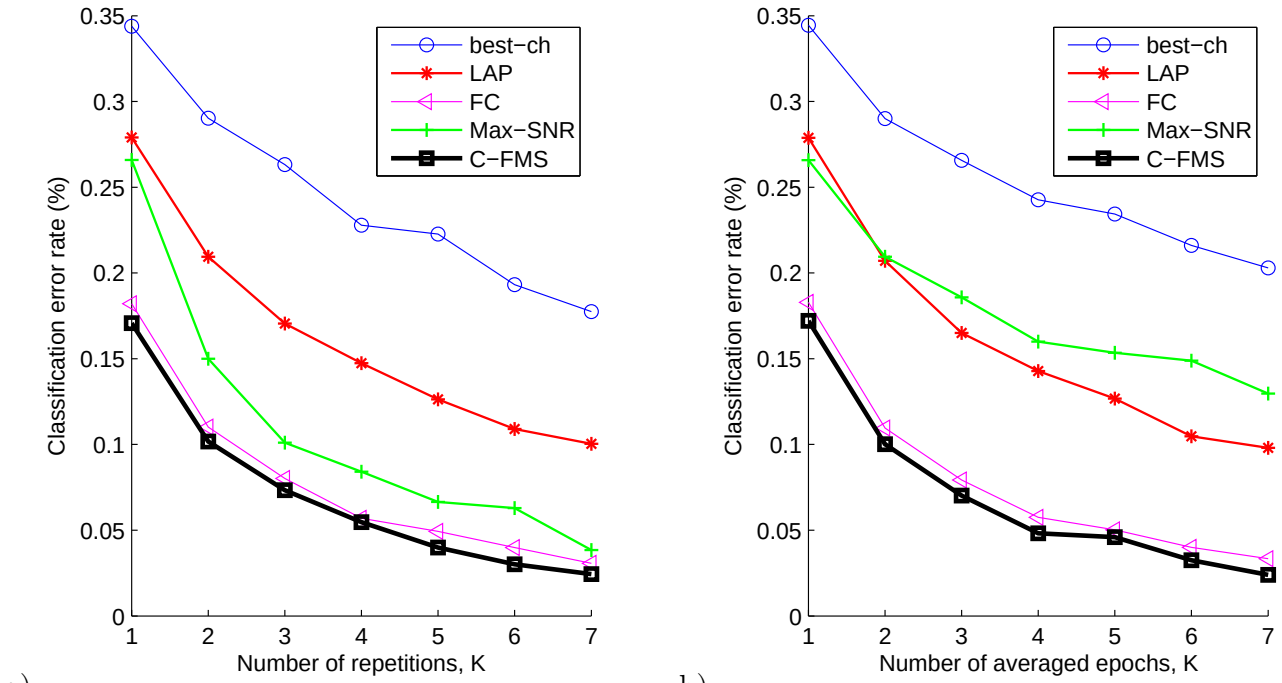


Figure 5.6: a) Classification results using the K-epoch average approach for $K \in \{1, \cdots, 7\}$. The results are the averaged values obtained from 23 subjects; b) Classification results using the K-probability approach for $K \in \{1, \cdots, 7\}$. The results are the averaged values obtained from 23 subjects.

i is computed by approximating the second derivative according to

$$V_i^{LAP} = V_i - \sum_{j \in S_i} g_{ij} V_j \tag{5.34}$$

where V_i is the potential between electrode i and the reference, and g_{ij} is given by

$$g_{ij} = \frac{\frac{1}{d_{ij}}}{\sum_{j \in S_i} \frac{1}{d_{ij}}} \tag{5.35}$$

where S_i is the set of electrodes j surrounding electrode i, and d_{ij} is the distance between electrodes i and j, $j \in S_i$. Laplacian derivations were computed at channels Cz and Pz, taking respectively (Fz, C3, C4, Pz) and (Cz, Oz, PO7, PO8) as surrounding electrodes.

Figure 5.6a) shows that the error rate decreases for all methods with increasing K, and a sharp decreasing is evident for $K \leq 3$. The C-FMS spatial filter has a smaller error rate than the remaining methods, for all K epochs. Table 5.3 shows the respective statistical

Table 5.3: Statistical tests (*t-test*) for the classification values obtained in Fig. 5.6a)

Number of epochs, K	**statistical tests for the error rate differences**
$K = 1$ (single epoch)	C-FMS - bst channel = 17.3% (t(22)=16.95, $p \leq 0.001$)
	C-FMS - Lap derivations = 10.8% (t(22)=12.17, $p \leq 0.001$)
	C-FMS - Max-SNR = 9.5% (t(22)=5.10, $p \leq 0.001$)
	C-FMS - FC = 1.1% (t(22)=3.65, $p \leq 0.001$)
$2 \leq K \leq 7$ (K-epoch average)	C-FMS - bst channel ($p \leq 0.001$) for $K \geq 2$)
	C-FMS - Lap derivations ($p \leq 0.001$) for $K \geq 2$)
	C-FMS - Max-SNR ($p \leq 0.001$, $K = 2$, $p \leq 0.005$, $K = 3$, $p \leq 0.05$, $K = 4$, $p \leq 0.05$, $K = 5$, $p = 0.084$, $K = 6$, $p \leq 0.005$, $K = 7$)
	C-FMS - FC ($p \leq 0.01$, $K = 2$, $p = 0.056$, $K = 3$, $p = 0.052$, $K = 4$, $p = 0.067$, $K = 5$, $p \leq 0.02$, $K = 6$, $p = 0.13$, $K = 7$)

tests. The error rate difference is almost always statistically significant, or approaches significance, except the difference between C-FMS and FC for $K = 7$.

In the K-probability approach, the NB classifier is applied to single epochs and the probabilities are combined using (5.33). Figure 5.6b) shows the classification results. The statistical t-test was again applied to evaluate the significance of the results. For a single epoch, the results are coincident with the K-epoch approach, since for $K = 1$, (5.33) is equal to (5.32). For $K = 2 \cdots 7$, the reduction of classification error rates between C-FMS, and best channel, Laplace derivations and FC, is very similar to the K-epoch average approach. The differences are statistically significant with $p \leq 0.001$ for best channel and Laplace derivations, and $p \leq 0.005$ for FC. The Max-SNR results are poorer than for the K-epoch average. The difference between C-FMS and Max-SNR is about 10% ($p \leq 0.001$). These results show that Max-SNR works better with data with higher SNR provided by the K-epoch average approach. The C-FMS filter is not affected because the feature selection algorithm selects mainly features from the FC filter.

5.4.4 Online results

In online operation, the binary classifier is applied to each one of the 12 events. Each event is classified as target or non-target with an associated *a posteriori* probability using (5.32) or (5.33). Online experiments used the K-epoch approach (5.32). The character recognition is now a 36-class classifier where *a posteriori* probabilities (scores) of the 12 events are combined to find the mentally selected character/symbol. The character is

found by detecting the respective row number, #row, and column number, #column. The following combination of the *a posteriori* probabilities was used

$$\boxed{\begin{array}{l}\text{if the number of events detected as target is} \geq 1, \quad \text{then} \\ \#_{row} = \underset{j\in\{1,\cdots,6\}}{\arg\max} P_{+}^{j} \ \wedge \ \#_{col} = \underset{l\in\{1,\cdots,6\}}{\arg\max} P_{+}^{l} \\ \text{else, if all events are detected as non-target, \quad then} \\ \#_{row} = \underset{j\in\{1,\cdots,6\}}{\arg\min} P_{-}^{j} \ \wedge \ \#_{col} = \underset{l\in\{1,\cdots,6\}}{\arg\min} P_{-}^{l}\end{array}} \tag{5.36}$$

where $P_{\{+,-\}}^{\{j,l\}}$ are the *a posteriori* probabilities associated with the events of rows (index j) and columns (index l). By words, if more than one event is detected as target, the method chooses the event most likely to be a target. If all the events are detected as non-target, then the method chooses the event less likely to be a non-target.

Each online session occurred after the respective calibration session. From the results of the calibration session, it was selected the least number of repetitions, K, for which an error rate up to 5-10% was found. The number of repetitions was then adjusted, when necessary, according to the online performance of the subject. The C-FMS was the selected spatial filter since it consistently provided better results during the pilot experiments and throughout the sessions in this study as confirmed by the offline analysis in the last section.

Under the same conditions that occurred during the calibration sessions, the subjects were asked to write a sentence. Subjects S1 to S12 (see Table 5.4) wrote the sentence 'THE-QUICK-BROWN-FOX-JUMPS-OVER-LAZY-DOG' (39 characters), subjects S13 to S19 wrote the sentence 'THE-QUICK-BROWN-FOX' (19 characters) and subjects S20 to S24 wrote the Portuguese sentence 'ESTOU-A-ESCREVER' (16 characters). Participants S13 to S24 wrote a shorter sentence since they underwent an additional paradigm during the same sessions (*Case-study II* described in chapter 6). The sentences were written at once without interruptions. In case of error, subjects could opt to correct the character using the 'del' symbol.

To assess the online performance and for comparison with state-of-the-art results, the number of decoded SPM, the ITR (3.3) and the PBR (3.8) were computed taking into account the ITI (3.5 s) and omitting the ITI. Table 5.4 summarizes these results, associated with number of repetitions (N_{Rep}) and online accuracy (3.9). The averaged results are presented for each group of subjects. From the group of participants with CP, subject S21 was unable to perform the online session because, during the calibration, the

algorithms did not detect target events with an accuracy above 80% even for $K \geq 7$, which was insufficient for online operation. The averaged results were obtained only from S20 and S22. Comparing the results of able-bodied and disabled participants, and taking the bandwidth as the main parameter, we see that on average the results are slightly lower for disabled participants. It is worth noting that almost all values were obtained for classification accuracies above 85%[2]. Comparing the online and offline results (Tables 5.4 and 5.5), we can observe that the online results match the expected results from the offline analysis. These results corroborate that the offine results may provide a good indicator of the online performance as long as the correct performance metrics are used.

5.4.5 Offline analysis

5.4.5.1 SNR and discrimination enhancement

One natural measure to evaluate the performance of the spatial filters is the SNR. It was estimated from (3.14) using all K epochs of calibration data sets. To assess the improvement performance, the SNRs of Max-SNR and FC beamformers were respectively compared with: 1) the SNR of the best channel; 2) the averaged SNR over the 12 channels; and 3) the SNR of Laplacian derivations at channels Cz and Pz, taking respectively (Fz, C3, C4, Pz) and (Cz, Oz, PO7, PO8) as surrounding electrodes. The SNR of C-FMS was not computed because its first projection coincides with the FC beamformer, and thereby would lead to the same results. The SNR estimates were then averaged taking 23 of the 24 subjects, achieving the results in Fig. 5.7. The results were obtained for different number of averaged epochs[3], $K = 1 \cdots 7$, thus simulating different numbers of repetitions of the events. The data sets from subject S21 were discarded in the analysis because this subject did not evoke a visible P300. The results are statistically evaluated with a t-test in Table 5.6. For all methods, as the number of epochs taken for average increases, the SNR also increases, which was expected given the phase-locked properties of ERPs. Comparing Fig. 5.6 and Fig. 5.7, a direct relationship between SNR and classification accuracy becomes apparent, i.e., methods with higher SNR provide a better classification. The exception goes to the Laplacian derivations, which shows a better classification than best channel and notwithstanding similar SNRs. The SNR improvements led to an

[2]Many of the participants wrote the sentences with a fewer number of repetitions (some of them with a single repetition) but with lower accuracies, so we chose not to show these results.

[3]It is important to note that, when averaging, the number of samples of the data sets is reduced by the number of epochs, K, used in the average. For instance, if $K = 2$ the number of target and non-targets epochs will be respectively $180/2 = 90$ and $900/2 = 450$; for $K = 3$, $180/3 = 60$ and $900/3 = 300$, and so on.

Table 5.4: *Case-study I*: online results.

			omitting ITI			with ITI		
Subject	P_{ac} (%)	N_{Rep}	SPM	ITR	PBR	SPM	ITR	PBR
S1	95.12	4	5.66	26.25	26.41	4.58	21.24	21.37
S2	95.12	5	4.62	21.41	21.53	3.87	17.95	18.06
S3	86.67	5	4.62	18.09	17.50	3.87	15.17	14.68
S4	95.12	3	7.32	33.94	34.14	5.61	26.01	26.16
S5	95.12	5	4.62	21.41	21.53	3.87	17.95	18.06
S6	86.67	5	4.62	18.09	17.50	3.87	15.17	14.68
S7	90.70	7	3.37	14.31	14.19	2.96	12.55	12.44
S8	79.59	5	4.62	15.66	14.12	3.87	13.13	11.84
S9	90.70	4	5.66	24.04	23.82	4.58	19.45	19.27
S10	90.70	5	4.62	19.60	19.42	3.87	16.44	16.29
S11	100.0	3	7.32	**37.83**	37.83	5.61	28.99	28.99
S12	86.67	4	5.66	22.19	21.46	4.58	17.95	17.37
Avg 1	91.02	4.6	5.22	22.73	22.45	4.26	18.50	18.27
S13	84.21	7	3.37	12.58	11.92	2.96	11.03	10.45
S14	82.60	5	4.62	16.66	15.56	3.87	13.98	13.05
S15	100.0	5	4.62	23.86	23.86	3.87	20.01	20.01
S16	90.47	4	5.66	23.93	23.69	4.58	19.36	19.17
S17	85.00	5	4.62	17.50	16.70	3.87	14.67	14.01
S18	95.00	2	10.34	**47.87**	48.13	7.23	33.45	33.64
S19	95.00	3	7.32	33.86	34.05	5.61	25.95	26.09
Avg 2	90.33	4.3	5.76	25.33	25.01	4.57	20.03	19.76
S20	100.0	6	3.90	**20.14**	20.14	3.35	17.33	17.33
S21	-	-	-	-	-	-		
S22	93.75	6	3.90	17.58	17.62	3.35	15.12	15.16
Avg 3	96.88	6	3.90	18.86	18.88	3.35	16.23	16.25
S23	100.0	5	4.62	**23.86**	23.86	3.87	20.01	20.01
S24	93.75	5	4.62	20.82	20.88	3.87	17.47	17.51
Avg 4	96.87	5	4.62	22.34	22.37	3.87	18.74	18.76
Overall Avg	91.82	4.7	5.23	23.11	22.86	4.24	18.71	18.51

Table 5.5: SPM and ITR (omitting ITI) using the offline classification accuracy in Fig. 5.6a) with C-FMS.

	Number of repetitions (K)						
	1	2	3	4	5	6	7
P_{ac}(%)	82.92	89.82	92.66	94.52	96.01	96.99	97.56
SPM	17.65	10.34	7.32	5.66	4.62	3.90	3.37
ITR	64.13	43.17	32.30	25.94	21.80	18.78	16.45

enhancement of the ERP and thereby to an increased discrimination between target vs.

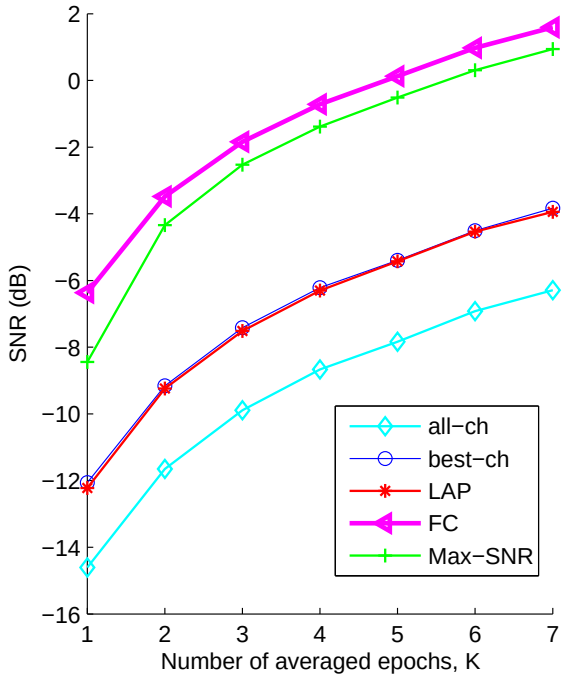
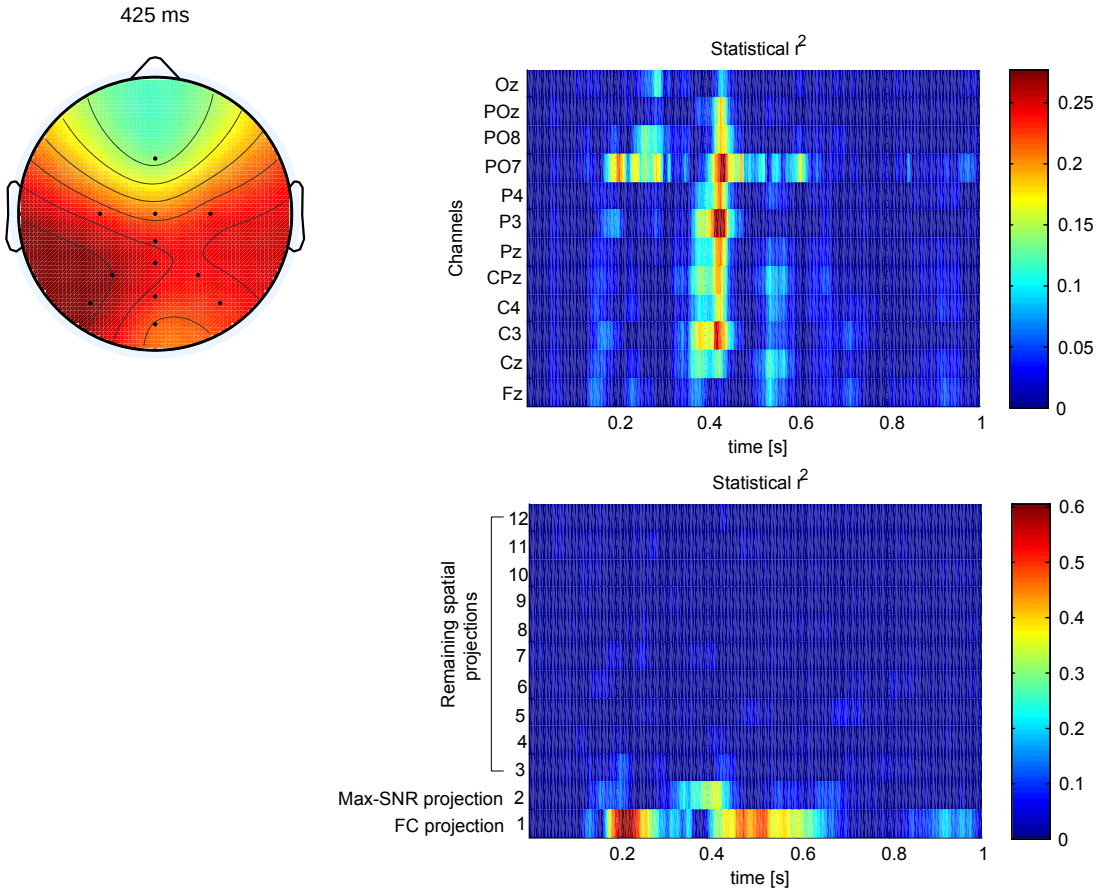


Figure 5.7: SNR estimated from 23 subjects. Analysis for K-epoch average, $K = 1 \cdots 7$.

Table 5.6: Statistical tests (*t-test*) for SNR values obtained in Fig. 5.7)

Number of epochs, K	**statistical tests for SNR differences**
$K = 1$ (single epoch)	FC - all-channel = 8.24 dB (t(22)=9.93, $p \leq 0.001$) FC - bst channel = 5.69 dB (t(22)=10.68, $p \leq 0.001$) FC - Lap derivations = 5.86 dB (t(22)=9.05, $p \leq 0.001$) FC - Max-SNR = 2.08 dB (t(22)=6.68, $p \leq 0.001$)
$2 \leq K \leq 7$ (K-epoch average)	Differences approximately constant: FC - all-channel = 8.18 dB; FC - bst channel = 5.68 dB; FC - Lap derivations = 5.76 dB; FC - Max-SNR = 0.87 dB; ($p \leq 0.001$ in all cases)

non-target. The statistical r-square measure was used to assess this discrimination. The color maps in Fig. 5.8 compares the r-square values before spatial filtering (top) and after C-FMS spatial filtering (bottom), for a representative data set with 180 target epochs and 900 non-target epochs. Channels with higher discrimination are usually over the parietal and parietal/occipital regions (typically, PO7 and PO8 provide the higher levels of discrimination). For C-FMS filtered data, the r-square was computed from projections (5.22). Projection 1 is the output of FC beamformer, $\mathbf{y_1}$, and projection 2 is the output of Max-SNR beamformer, $\mathbf{z_1}$. The remaining projections are obtained



a) b)

Figure 5.8: Results obtained from representative data of one session: 180 target epochs and 900 non-target epochs using a 5-epoch average. Color map representing the r-square statistical measure of the discrimination between target and non-target classes. a) Topographic map before spatial filtering at instant 425 ms; b) Top: r-square at channels Fz, Cz, C3, C4, CPz, Pz, P3, P4, PO7, PO8, POz and Oz before spatial filtering ($\mathbf{X}_+, \mathbf{X}_-$); b) Bottom: r-square of projections of C-FMS beamformer according to (5.22), where projection 1 is the output of FC beamformer, $\mathbf{y_1}$, and projection 2 is the output of Max-SNR beamformer, $\mathbf{z_1}$. The remaining projections are $\mathbf{Z}^{(2:N-1)}$ according to (5.20).

from $\mathbf{Z}^{(2:N-1)}$ (5.20). As expected, the first C-FMS projection shows the higher r-square discrimination, increasing the pre-filter maximum of approximately 0.3 to a 0.6 post-filter maximum. Although lower, the second projection of C-FMS also shows some degree of discrimination. The other projections show no discrimination. This result confirms that FC and Max-SNR outputs retain the most discriminative information. Figure 5.9 shows the mean, $\mu(t)$, and mean $\pm$ standard deviations, $\mu(t) \pm \sigma(t)$, of target and non-target epochs measured at each instant t at channel Cz before spatial filtering (top), and $\mu(t)$ and $\mu(t) \pm \sigma(t)$ of first C-FMS projection (bottom). The increased margin of separation

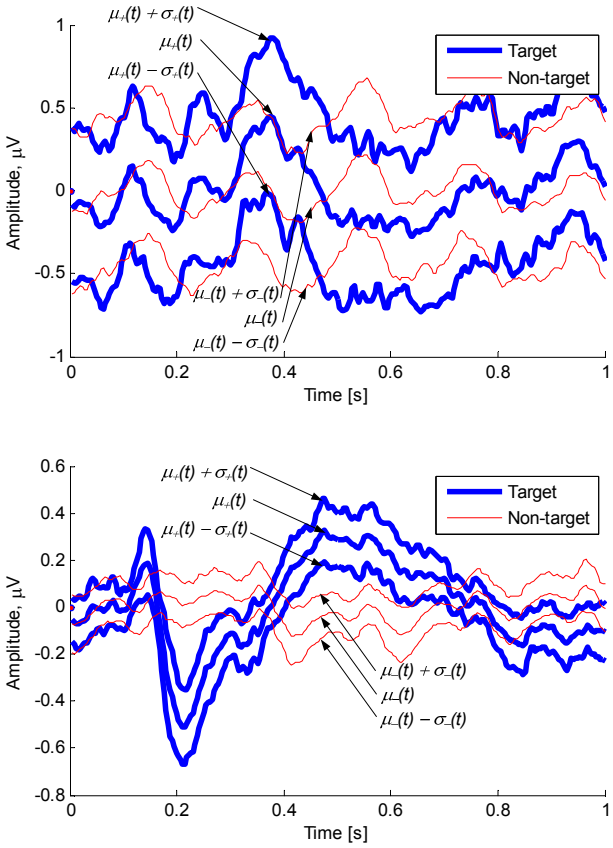


Figure 5.9: Results obtained from the same data shown in Fig. 5.8. Top: mean, $\mu(t)$, and mean $\pm$ standard deviation, $\mu(t) \pm \sigma(t)$, of 180 target epochs and 900 non-target epochs measured at channel Cz using a 5-epoch average; Bottom: $\mu(t)$ and $\mu(t) \pm \sigma(t)$ of the first C-FMS projection, $\mathbf{y_1}$, of the 5-epoch average of 180 target epochs and 900 non-target epochs.

between the patterns of the two classes after C-FMS filtering is remarkable. Figure 5.10 shows also the effect of spatial filtering in the frequency domain. The plot represents the average of the FFTs spectra of the first spatial projection. Comparing with Fig. 5.3(d), it can be seen that the 5 Hz interference is almost eliminated from target epochs.

5.4.6 Spatial filtering robustness

The importance of spatial filtering and its advantage over the classical 'spatio-temporal' approach is further analysed. Target ERPs exhibit an inter-trial variability regarding latency, amplitude and morphology as can be seen in example of Fig. 5.11. The plots compare 10 successive epochs before (a) and after (b) spatial filtering, emphasizing epoch variability and effect of spatial filtering. Despite this inter-trial variability, there is a

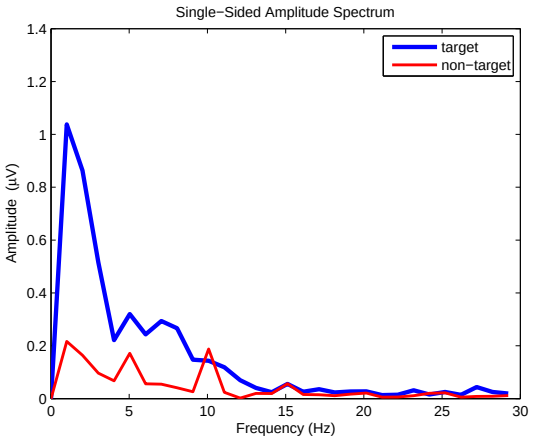


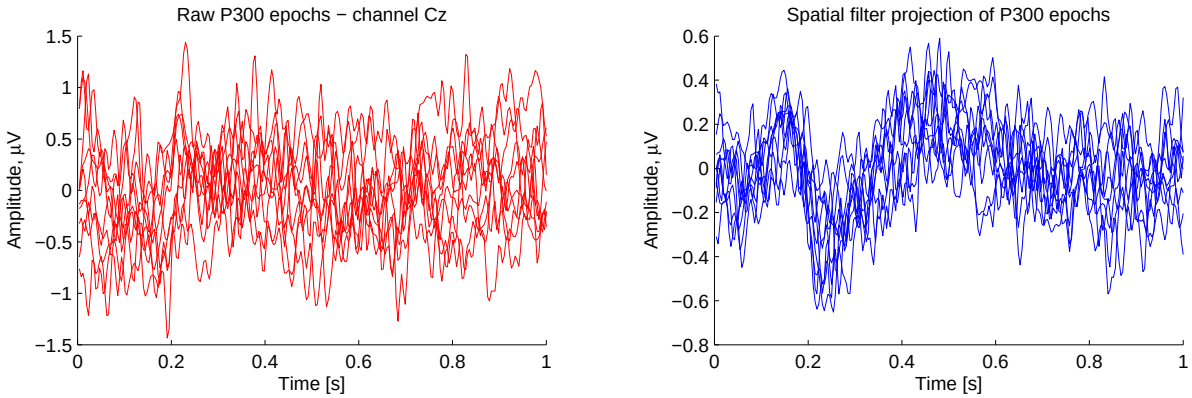
Figure 5.10: Average of the FFT spectra of the first projection obtained from C-FMS filtered epochs (180 FFTs of target epochs and 900 FFTs of non-target epochs).

spatial correlation between channels, that we will show to be invariant across trials in normal conditions. Spatial filtering takes advantage of this neurophysiological invariance on spatial correlation, giving it a robustness to inter-trial variability. Let us ascertain the spatial correlation through a coherence measure. It gives a linear correlation between two signals as a function of the frequency. In the neurophysiologic context, coherence can be used to measure the linear dependence and functional interaction between different brain regions. The coherence between two signals x and y was estimated from the magnitude squared coherence [Bendat 2000]

$$\begin{aligned}\lambda^2_{xy}(k) &= \frac{|S_{xy}(k)|^2}{|S_{xx}(k)||S_{yy}(k)|} \\ &= \frac{|X(k).Y^*(k)|^2}{|X^2(k)||Y^2(k)|} \\ &= \frac{||X(k)|.|Y(k)|e^{j(arg(X(k)-arg(Y(k))}|^2}{|X^2(k)||Y^2(k)|}\end{aligned} \tag{5.37}$$

where * denotes the complex conjugate, $S_{xx}(k)$ and $S_{yy}(k)$ are respectively the power spectra of $x(n)$ and $y(n)$, and $S_{xy}(k)$ is the cross-power spectrum. The coherence is sensitive to both phase and amplitude of the relationship between pairwise channel signals, as can be seen in (5.37), and thus provides a reliable measure. Fig. 5.12 shows the spatial correlation between CPz and several other channels (Cz, Pz, P3, P4, PO7 and PO8). For each trial, the coherence values (5.37) were averaged for a frequency band according to

$$\overline{\lambda^2_{xy}} = \sum_k \lambda^2_{xy}(k) \tag{5.38}$$



a) b)

Figure 5.11: a) Plot of 10 successive unfiltered 5-epochs average (P300 epochs at channel Cz) of a representative participant. The variability of amplitude, latency and morphology make very challenging the detection of an ERP pattern, even with a 5-epochs average; b) Spatial filter projections, obtained from 12 channels, for 10 successive 5-epochs average (using topographic data of the same temporal epochs on a)).

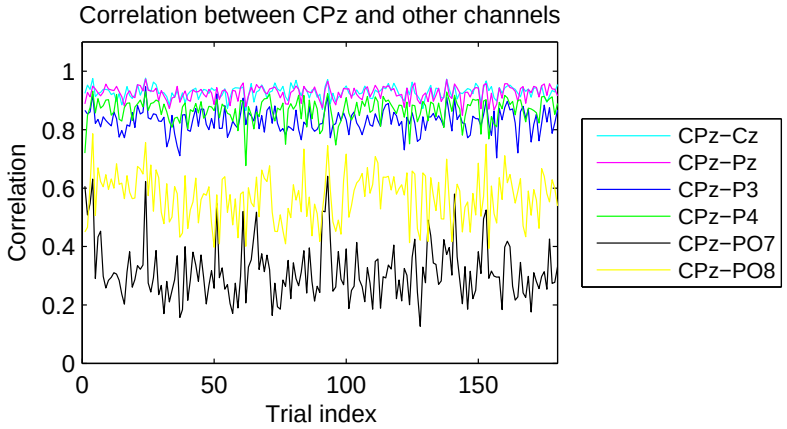


Figure 5.12: Spatial correlation measured between channel CPz and several other channels using (5.38) across 180 target epochs. The correlation remains invariant even for channels far from CPz where the correlation is smaller.

where the summation goes though all the frequencies of a given frequency band. The invariance of the correlation across 180 target epochs is particularly evident in the near proximity but also in far locations such as PO7 and PO8.

To test the robustness of spatial filtering, we compared the FC beamformer estimated from two independent data sets obtained from the same subjects gathered in different moments of the experimental sessions. Figure 5.13 compares the first spatial filter w_1 obtained from each of the datasets. The analysis was made for one subject of the able-

bodied group, one subject of the CP group and one subject of the ALS group. Fig. 5.13 presents at top the weights of the spatial filters while at the bottom a colormap represents the interpolation of the same filter coefficients at electrode positions. The two filters obtained from the two data sets are very similar, which shows that the spatial filters have a good generalization (with no training overfit), providing invariant features for the classifier.

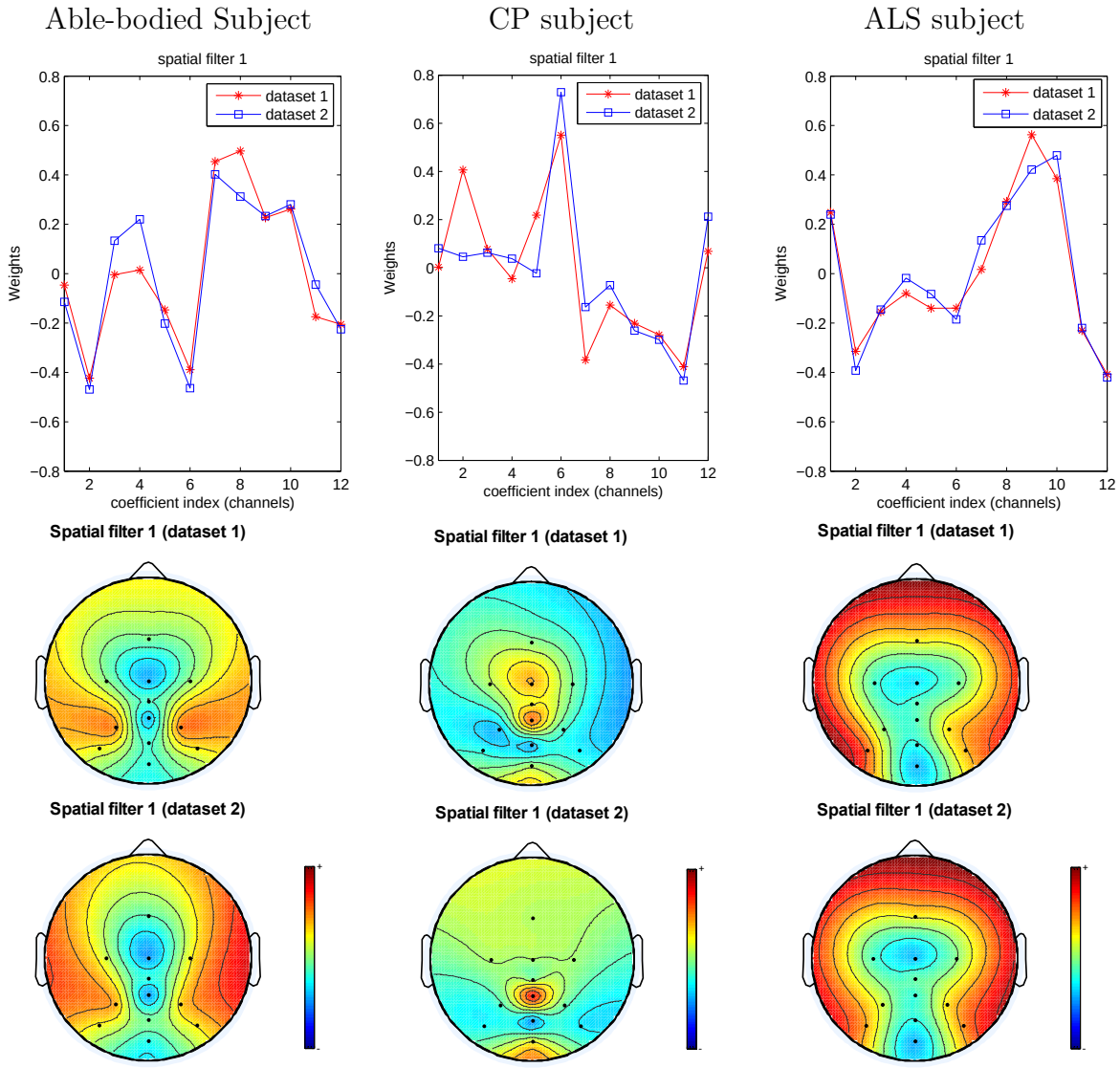


Figure 5.13: Analysis of spatial filter robustness. Top: the graph compares the coefficients w_1 of the spatial filters estimated from two different datasets (dataset 1 and dataset 2) for the same subjects. The analysis was made for one able-bodied subject (left), one subject of the CP group (middle), and one subject of the ALS group (right); Bottom: scalp topography of the same spatial filters shown in top. The colormap represents the interpolation of coefficients w_1 at electrode positions.

5.5 Assessment of spatial vs. spatio-temporal classification

As referred in section 5.1.1, most of the P300-based systems follow a 'spatio-temporal' approach (as illustrated in Fig. 5.1). From a machine learning perspective, our spatial filtering approach has comparatively several advantages, some of them already inferred previously:

- Spatial projections have a reduced inter-trial variability, providing robust features to classifiers. The robustness of classification models reduces the overfitting problem. A 2-dimensional example is illustrated in Fig. 5.14. The scatter plots show the two best features selected from the 12 channels, using a spatio-temporal approach (left), and the two best features selected from the spatial projections, using a spatial approach (right). Class separation is notoriously higher in the spatial approach;

- It transforms the original spatial feature space into a new one with higher SNR that concentrates all the energy, eliminating the spatial redundancy;

- The dimension of the feature space is substantially reduced ($N \times T$ dimension is transformed into a $1 \times T$ dimension or $2 \times T$ dimension). Thereby, hyper-parameters of classifiers are tuned more quickly and more efficiently, and regularization may be unnecessary for some classifiers. Moreover, the number of training samples can be reduced, requiring shorter calibration periods;

- Robust features in a low dimensional feature space make classification less dependent of both feature selector type and classifier type.

Additionally, the criteria used to estimate spatial filters are based on physical properties of the signals, such as power and crosspower (physically meaningful), which are computed using the full time series. On the other hand, spatio-temporal classification often disregards the spatial correlation and physical properties of the signals, discarding valuable information. Moreover, due to high dimensionality of the feature space, $N \times T$, a downsampling operation is required, otherwise overfitting problems occur and mathematical models may become intractable or unsolvable. The next section compares the two approaches using benchmarking datasets.

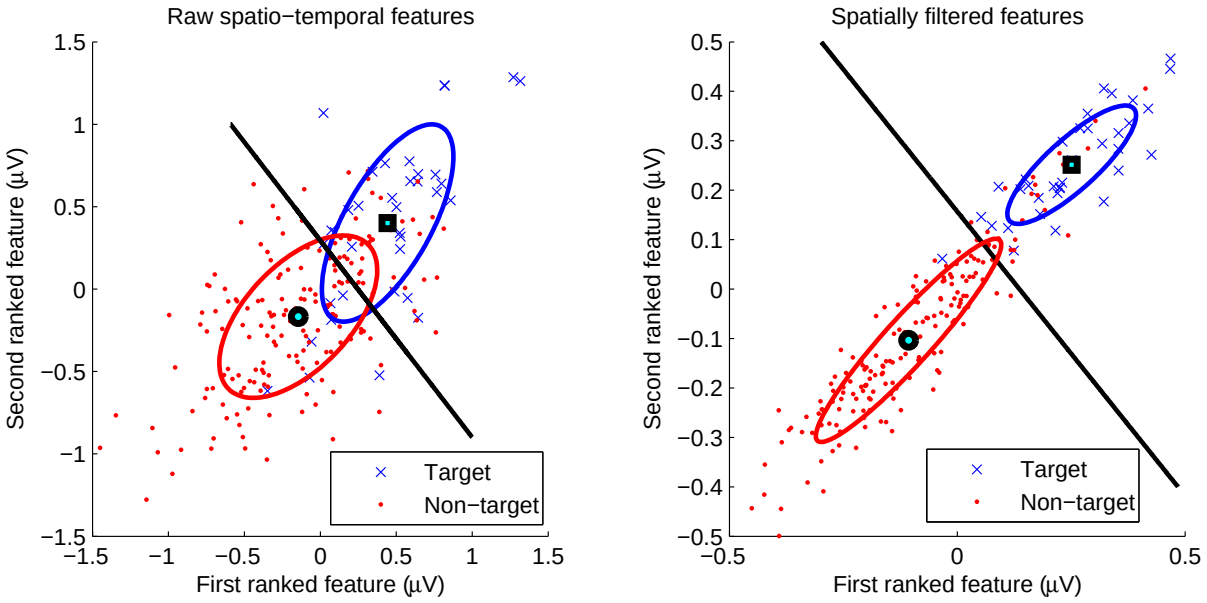


Figure 5.14: 2D scatter plots for the two best features ranked with r^2. Features were obtained from 5-epoch average using an RC speller dataset. Covariance contour and mean is shown for each class. Separating hyperplane computed using LDA, considering equal covariances; Left) Features selected from the 12 channels, using a spatio-temporal approach; Right) Features selected from the spatial projections, using a spatial approach. Note the larger inter-class distance for spatially filtered features.

5.5.1 Benchmarking datasets - BCI Competition II

The goal of this section is twofold. First, to infer the impact of spatial filtering compared to spatio-temporal approach, independently of the classifier type. Therefore, we compared the two approaches by using several standard classifiers presented in chapter 4: LDA (section 4.4.1), SWLDA (section 4.4.2), qNB (section 4.4.3), SVM (section 4.4.4). Note that, although an implicit comparison between classifiers is made, our goal is beyond this comparison, aiming to show the effective contribution of spatial filtering for any of the classifiers. Second, to compare the performance of our proposed methodology with state-of-the-art methods.

The BCI-Competition 2003 [BCI-Competition 2003] made available datasets gathered during an RC speller task performed by one participant. During the task, the participant had to spell the following words 'FOOD-MOOT-HAM-PIE-CAKE-TUNA-ZYGOT-4567' (31 characters). Simulating the conditions of the competition, we obtained the classification models from labeled data sets (sessions 10 and 11), which were then tested on unlabeled data sets (session 12). See details of the datasets in [BCI-Competition 2003].

Table 5.7: Number of errors in the recognition of 'FOOD-MOOT-HAM-PIE-CAKE-TUNA-ZYGOT-4567' (31 characters) obtained with several classifiers using either a spatio-temporal or a spatial approach.

	Number of event repetition K				
Method	1	2	3	4	5
LDA (spatial)	21	10	5	2	0
LDA (spatio-temporal)	27	18	20	9	6
SWLDA (spatial)	18	13	7	2	4
SWLDA (spatio-temporal)	25	15	12	9	7
qNB (spatial)	17	12	9	0	2
qNB (spatio-temporal)	29	28	25	26	25
SVM (spatial)	21	14	6	1	0
SVM (spatio-temporal)	27	23	17	20	14

Following the spatio-temporal and spatial approaches, the classifiers were tested achieving the results in Table 5.7. In the spatio-temporal approach, the EEG from each channel (the same 12 channels used in our own experiments) was downsampled with appropriate moving average filtering by a factor of 10, and then normalized and concatenated into a unidimensional feature vector. The 100 best features were then selected using the r^2 feature selector (4.2), except in the SWLDA case, which used its embedded feature selector. The criteria to enter and remove features was respectively $p - value < 0.1$ and $p - value > 0.15$. The SVM used an RBF Kernel, and parameters C and σ were tuned through grid-search and assessed by cross-validation. The classifier models were achieved for different number of event repetition, $k \in \{1 \cdots 5\}$ considering for each case the average of k epochs. In the spatial approach, the 12 channels were spatially filtered through the C-FMS method, and the 2 projections were concatenated into a unidimensional feature vector. No downsampling was performed. The remaining steps were similar to the spatio-temporal approach. The 36-class character recognition was then performed using (5.36) (the chance level is 2.7 %).

For all classifiers, it is notoriously the improvement of performance for the spatial approach. While all classifiers perform well with the spatial approach, the results are poor for the spatio-temporal approach. The worst cases with the spatio-temporal approach are with qNB and SVM. In the spatial approach, the simplest classifier, qNB, has a similar performance to the other methods, even outperforming for some number of repetitions, k. Actually, it achieves the best performance for $k = 4$. The inferred words and error rates for this combination are shown at Table 5.8. These results outperform the results achieved by the winners in [BCI-Competition 2003]. No individual adjustments of the number of features and training sets was made to each classifier, and we limited the number of

Table 5.8: Inferred words and associated error rates for different number of repetitions, using data sets from BCI - Competition 2003, using (spatial + qNB) approach.

N_{Rep}	Inferred words	Error
1	FCOD MMOZ NBM JID CACC TTNC ZZBUT XXT7	54.8 %
2	FCOD GMOT BAM JIE CAIE TCNA ZMAOT X0Z7	38.7 %
3	FOOD MOOT BAM JIE CAKC TCNA ZSAOT X057	29.0 %
4	FOOD MOOT HAM PIE CAKE TUNA ZYGOT 4567	0 %

channels to 12. Therefore it can be expected a further performance improvement of the results.

5.6 Conclusion

This chapter described several novel statistical spatial filters for ERP classification. The performance regarding neurophysiologic and classification results are compelling, showing that spatial approach should be considered as an effective option in current ERP-based BCIs. Although the methods were applied in a P300-based BCI framework, they can also be used to reduce the recording duration in patient examinations, when P300 detection is used as a diagnostic tool (e.g., cognitive impairments, neurological and psychiatric disorders) [Mell 2008], or as described in the clinical application of appendix D.

Chapter 6

Lateral Single Character (LSC) Speller

Contents

"Intelligence is the ability to adapt to change."
"Remember to look up at the stars and not down at your feet."
Stephen Hawking, one of the brightest physicists of today (Stephen Hawking suffers from an atypical amyotrophic lateral sclerosis since the age of 21)

This chapter describes a novel speller paradigm called lateral single character (LSC) speller. This paradigm follows an event strategy that significantly reduces the time for symbol selection and explores the intrinsic hemispheric asymmetries in visual perception, to improve the BCI performance. The LSC performance is compared to that of the standard row-column (RC) speller, through tests performed by able-bodied and individuals with severe motor limitations (*Case-study II*).

6.1 Introduction

P300-based BCIs should be designed to enhance simultaneously: (1) the 'quality' of the P300 component, which has direct influence on the classification accuracy; (2) the amount of encoded information; and (3) the number of decoded symbols per minute. Increasing one of these parameters while decreasing the others may have little impact on the improvement of communication rates. It is known that the P300 ERP represents cognitive functions that vary according to stimulus modality, stimulus probability, stimulus onset asynchrony (SOA), task-relevance, decision making, selective attention, expectancies and relative perceptual distinctiveness among stimuli [Polich 2007, Rohrbaugh 1974, Polich 1996, Comerchero 1999, Heinrich 2008]. In essence, a P300-based BCI uses an attention task where temporal, spatial, object-based and featural attention can be manipulated [Correa 2006, Mercure 2008]. Thereby, there are many parameters that can modulate the P300 component and that can be configured in P300-BCIs to enhance the evoked P300. On the other hand, issues such as user-friendliness and

comfort, should not be overlooked when the interface is intended to be used on a daily basis.

6.2 Design of visual paradigms

6.2.1 Limitations of the standard RC speller

The original RC speller is the most widely used P300 speller and has already shown to achieve effective transfer rates for clinical use [Kubler 2008]. It encodes a large number of symbols with a small number of events, which potentiates high information transfer rates (ITRs). Notwithstanding, the RC speller presents some issues that limit its performance. Two known problems identified from our own experience and well documented in [Townsend 2010] are the adjacency-distraction errors and the double-flash errors. In the first case, non-targets rows/columns adjacent to target rows/columns may distract or deceive the user and be erroneously selected. This is an example of spill-over of spatial attention. The second case occurs when there are two consecutive target flashes. The second flash is usually unnoticed by the user, and does not evoke a P300 signal. Even if the second flash is noticed, there will be an overlap of the two components causing a change of the P300 waveform morphology increasing its variability [Martens 2009]. This is an example of spill-over of temporal attention and target-to-target interactions. The frequent occurrence of overlap in the RC paradigm is due to its short target-to-target interval (TTI), which is in turn related with the short SOA (stimulus onset asynchrony - interval between the onset of two consecutive stimuli) and with the high target probability (1:6). The P300 amplitude is adversely affected by the high target probability because amplitude and probability are inversely proportional. Overcoming these two aspects by modifying the original RC paradigm represents a considerable challenge. For example, increasing the SOA leads to a lower ITR. Moreover, a decrease of the probability can only be achieved with a larger matrix, possibly encoding unnecessary symbols and increasing the time for symbol selection.

6.2.2 New visual paradigms and RC modifications

Several studies have investigated new strategies for stimuli presentation and the impact of their physical properties (see a summary in Table 6.1). Studies in [Allison 2003b, Sellers 2006b, Nam 2009] investigated the effect of changing the size of the original RC matrix on the P300 amplitude, and showed that larger matrix sizes increased the am-

plitude. Notwithstanding, in [Sellers 2006b] the smaller matrix reached higher accuracy, possibly due to the better separated spatial attention signals (better attentional 'zoom-in'). The effect of different SOAs in the RC speller was also analyzed and non-consensual results were reached. In [Farwell 1988, Nam 2009], a longer SOA increased the P300 classification accuracy. In [Sellers 2006b, Meinicke 2002], a shorter SOA led to a higher accuracy, but to lower P300 amplitudes [Sellers 2006b]. In experiments addressing either probability or the role of SOA, it seems that there is no straightforward probability and SOA relation between P300 amplitude and classification accuracy. This may be due to the fact that one may have to take into account the joint effects of spatial, temporal and object based attention, and the final result will depend on an amplitude vs. overlap tradeoff.

P300 is considered an endogenous component not very influenced by the physical attributes of the stimuli [Sanei 2007]. However, such physical attributes might indirectly affect the response by modulating perceptual saliency as suggested by [Teixeira 2010]. Accordingly, exogenous attributes such as intensity, color and size can enhance stimulus discrimination and perception, eliciting higher and faster components [Polich 1996, Treder 2010]. In [Salvaris 2009] and [Takano 2009], several physical parameters of the original RC speller were changed and assessed, namely, background and symbol color, symbol size, inter-symbol distance, and luminance/chromatic flicker.

New flash-patterns presentations have been proposed as alternatives to the standard RC approach. In [Allison 2006] a multi-flash approach was analyzed aiming to reduce the number of necessary flashes within a trial to detect a symbol. A fewer number of flashes can lead to an increased ITR if the level of accuracy is kept. [Townsend 2010] introduced a new paradigm referred to as checkerboard that relies on a "splotch" stimulus presentation [Allison 2003a], where the stimuli of the matrix are presented in predefined groups instead of rows and columns. The checkerboard approach eliminates the double flash effect, reduces the problem of component overlapping, and avoids the adjacency-distraction issue. Despite the fact that the standard matrix is increased to a 8×9 matrix requiring 24 flashes per selection, thus increasing the time to make a selection, the achieved results show even so an increased performance over the standard RC paradigm. Also based on the "splotch" concept, the use of a different number of flashes was compared, showing improvements over the RC speller [Jin 2011a].

An alternative to the RC-based speller is the single-character (SC) speller wherein each symbol is individually highlighted. In the SC speller, each event only encodes a single symbol, resulting in low ITR. Moreover, the number of distractor symbols is larger than

Table 6.1: Summary of new stimuli presentation paradigms and the effects of stimuli properties.

Stimuli properties and presentations	Studies
Manipulation of RC properties:	
- Effects of matrix size and ISI	[Allison 2003b, Sellers 2006b, Nam 2009]
- Effects of SOA	[Farwell 1988, Meinicke 2002, Sellers 2006b, Nam 2009]
- Effects of physical attributes: color, luminance, size, inter-symbol distance	[Salvaris 2009]; [Takano 2009]
Manipulation of flash patterns:	
- Multi-flash	[Allison 2006]
- 'Splotch'	[Allison 2003a, Townsend 2010, Jin 2011a]
- Single-char	[Guan 2004]; [Guger 2009]
Gaze independent paradigms:	
- Two-level based	[Treder 2010]; [Pires 2011c]
- Sequential central disk	[Liu 2011]

in RC, since the number of flashes is the same as the number of symbols. However, on the other hand, the SC speller has a lower target probability and the TTI is larger, which are two parameters that can enhance the P300 amplitude and avoid overlapping. [Guan 2004] compared the SC speller with the RC speller and achieved significantly better results with the SC speller. However, in a study made by [Guger 2009] with 100 participants, the results showed a significantly higher accuracy with the RC speller. Although the P300 amplitude has been higher for the SC speller than for the RC speller, the average accuracy was worse for SC. Several speller paradigms addressing the gaze-independence issue have also been proposed. This theme will be discussed in detail in chapter 7.

6.3 Novel paradigm: Lateral single-character speller

This study proposes a novel speller, henceforth called lateral single character (LSC) speller (described in section 6.3.1). It introduces new features that overcome some of the limitations found in current SC spellers and in standard RC spellers. Namely, compared to other SC spellers, the layout reduces the effect of local and remote distractors, and the flashing/event strategy allows to drastically reduce the SOA. This reduction makes the overall time of one trial very similar to those usually achieved in the RC speller with

the same amount of encoded symbols. By bringing together these features with the inherent lower target probability and higher TTI of SC spellers, LSC elicits a P300 with higher 'quality', thus improving the classification accuracy. Furthermore, the paradigm is expected to be more visually attractive and comfortable.

In this study, LSC and RC spellers are compared based on online sessions performed by able-bodied and disabled persons, by assessing the accuracy, symbols per minute and ITR. Participants answered to a questionnaire to assess both subjective and objective parameters of the interfaces. The BCI performance is also compared with the performance achieved with non-EEG interfaces for those participants who use standard interfaces in their daily lives. Datasets gathered during calibration sessions were used for an offline analysis. We compared the amplitude, latency and signal-to-noise ratio (SNR) of the ERPs elicited by the RC and LSC paradigms. Regarding LSC, new discriminative neurophysiologic features related with the lateral layout and event strategy are investigated and discussed in section 6.6.3.

6.3.1 LSC design

The proposed LSC speller is shown in Fig.6.1a). It encodes 28 symbols comprising all letters of the alphabet, and the 'spc' and 'del' symbols. It has a lateral and symmetrical arrangement where the symbols flash alternately between the left and right fields of the screen (see the temporal diagram of Fig. 6.1c)). There is always a symbol highlighted, thus the ISI (inter-symbol interval) is eliminated. The SOA coincides with the highlight time of a symbol, which substantially reduces the sub-trial time. The symbols are flashed pseudo-randomly since the side of consecutive flashes is controlled, but within each side the symbols are flashed randomly. The user is overtly focused on the left or on the right side of the screen and therefore the participant sees virtually only half of the stimuli, because stimuli on the opposite side are nearly ignored (remote distractors). Likewise, despite the target probability being 1:28, the user perceives virtually a target probability of 1:14. Moreover, the user sees the on-off visual effect although there is always an active symbol, i.e., the ISI is zero, but for each side of the screen there is respectively a left ISI and a right ISI. In the experiments, the SOA was settled to 75 ms and thus each round of flashes took $28 \times 0.075 = 2.1$ s. The symbols are arranged to minimize the effect of local distractors. This effect depends on the number and distance of the surrounding stimuli. In the typical 6×6 RC or on a SC matrix layout, a central symbol has 8 surrounding distractors, while in the LSC speller the maximum number of surrounding distractors is 4 (see Fig. 3.1d)). The circular layout of the speller provides similar eccentricities for all

symbols, avoiding large eye movements to see symbols in the corners, as occurs in RC. The highlighting of each symbol is made by changing the foreground/background color of the symbol from white/grey to red/green, and by increasing the size by $\approx 15\%$. Finally, in the LSC speller, the spelled or copied letters are presented in a central position of the screen, while in the RC speller they are presented at the top of the screen. The central position avoids large eye movements to check the detected spelled symbol. It should be emphasized that despite the control of P300-based ERPs requires overt attention, the movement of the eyes does not need to be precise as in the case of SSVEP. An eye movement placing the target event in the foveal region is enough for its detection.

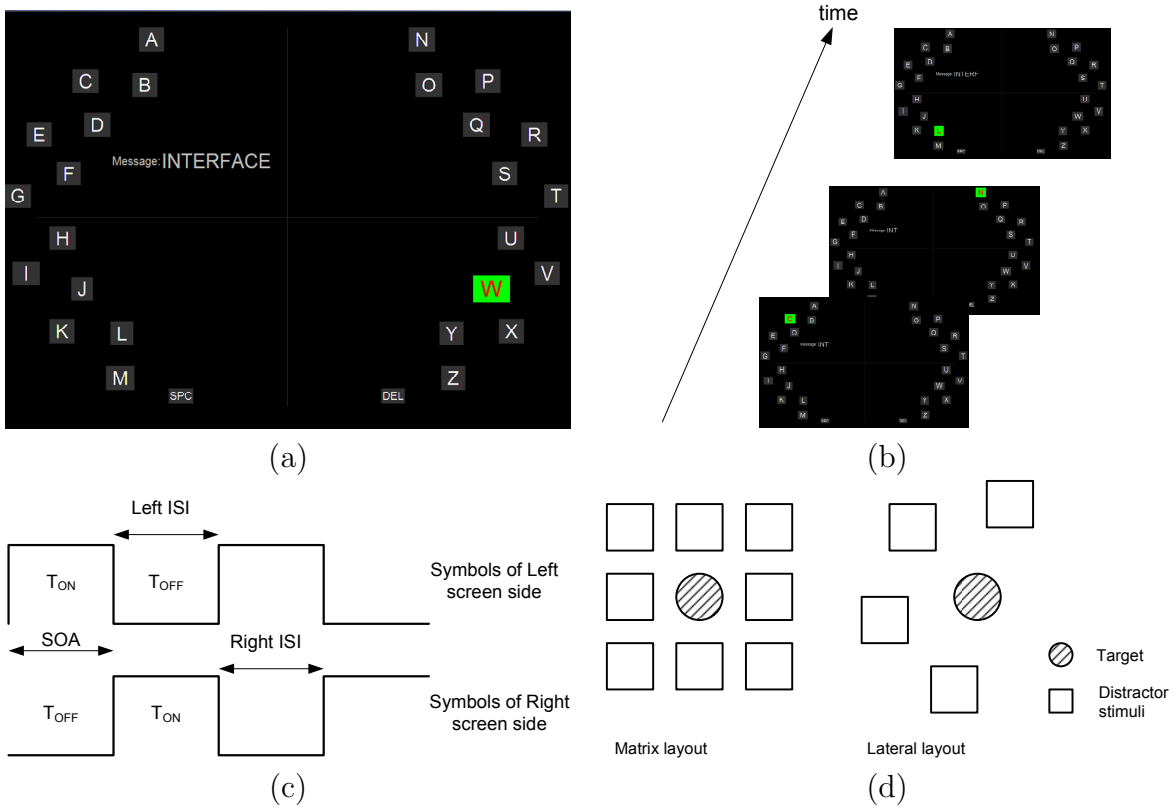


Figure 6.1: a) Screenshot of the LSC speller; b) Screenshots of successive events; c) Temporal diagram of the LSC stimuli/events. T_{ON} is the highlight time and T_{OFF} is the time between flashes on each screen side (Left ISI or Right ISI). T_{ON} has to be equal to T_{OFF}; d) Local distractors on a matrix layout and on the LSC layout.

6.4 *Case-Study II*: LSC vs. RC

6.4.1 Participants

The experiments were performed by ten able-bodied participants, five participants with cerebral palsy (CP), one participant with Duchenne muscular dystrophy (DMD), one participant with spinal cord injury (SCI) and seven participants with amyotrophic lateral sclerosis (ALS). Table 6.2 shows a summary of clinical data of the participants with motor disabilities, as well as their main signs, levels of functionality, and interfaces currently used in their daily lives. Participants S11, S12, S18, S19 and S20 already participated in *Case-study I*. Participants S18-S21 and S24 are confined to a wheelchair, and use adapted interfaces to steer the wheelchair and to access the computer. Participant S23 is also confined to a wheelchair but cannot control it, and recently lost the ability to control an head-tracker. All these participants suffer from severe motor disabilities and are highly dependent of human assistance. Participant S22 can walk but requires a special device to access the computer keyboard. One of the above individuals is in the borderline range of intellectual functioning. Participants S11-S17 have ALS with either bulbar or spinal onset. Their main signs are respectively dysarthria and dysphagia, and muscular weakness. None of the ALS individuals are in an advanced stage of the disease and therefore none of them currently uses adapted interfaces. The group of able-bodied participants is composed of 3 males and 7 females with ages ranging from 24 to 38 years old, averaging 29.2 years old. Able-bodied participants are referred to as group I, ALS participants are referred to as group II, and for sake of concise presentation, CP, DMD and SCI participants, all suffering from severe communication and mobility limitations, are referred to as group III. Participants of group III, all use in their daily lives non-EEG interfaces to control the computer and/or the wheelchair, and thus their BCI performances were compared to those achieved with their usual non-EEG interfaces.

6.4.2 Calibration and online sessions

The experiments consisted of calibration and online sessions that took place at facilities of the Cerebral Palsy Association of Coimbra (APCC) (group III), at the Hospitals of the University of Coimbra (HUC) (Group II) and at our laboratory facilities (Group I). The online sessions were preceded by a calibration session. The participants were instructed to be relaxed and attend to the desired target, while mentally counting the number of 'perceptual intensifications' of such target events. Participants S18-S21, S23-S24 and S14

Table 6.2: Motor disabled individuals: summary of main signs, functional status, and interfaces.

Subject/ Sex	Age	Diagnosis	Time since diagnosis	Main signs / Main functionality	Adapted interfaces and assistive devices
S18/F	18	CP	Posnatal	Tetraparesis, dystonia with spasticity and dysarthria / Head control	Powered wheelchair: head-switch connected to a scanning interface; Computer: head-switch and head-tracking
S19/M	34	CP	Perinatal	Tetraparesis, dystonia with spasticity, and high dysarthria / Head and right foot control	Powered wheelchair: adapted joystick controlled by right foot; Computer: controls mouse and keyboard with the right foot
S20/M	46	CP and discal hernia C3-C4	Perinatal	Tetraparesis, spasticity and dysarthria / Head control	Powered wheelchair: joystick adapted to chin; Computer: head-tracking
S21/M	45	CP	Neonatal	Tetraparesis, choreoathetosis and high dysarthria / Head and feet control	Manual wheelchair: foot; Computer: helmet with pointing device to select letters of the keyboard
S22/F ID07	42	CP	Neonatal	Tetraparesis, ataxia (high motor incoordination) and dysarthria / Walking ability and coarse control	Computer: grid over the keyboard and use of track-ball (with difficulty)
S23/M ID05	30	DMD	>22 years	Tetraplegia / Slight movements of the head, jaw control	Manual wheelchair: no control; Computer: used to control head-tracking
S24/M ID06	28	Spinal cord injury: C3-C4 lesion	12 years	Complete spatic tetraplegia and tracheotomy / Slight movements of the head and of the left upper limb	Powered wheelchair: joystick adapted to chin; Computer: head-tracking
S11/F	67	Bulbar-onset ALS (FRS-r 46)	7 years	Dysarthria, dysphagia, and muscular weakness in upper limbs / -	-
S12/F	75	Bulbar-onset ALS (FRS-r 40)	1 year	High dysarthria and dysphagia / -	-
S13/M	58	Bulbar-onset ALS (FRS-r 47)	1 month	Slight dysarthria and dysphagia / -	-
S14/F	78	Spinal-onset ALS (FRS-r 32)	1 month	Muscular weakness in limbs and high dysarthria / -	Manual wheelchair
S15/M	80	Bulbar-onset ALS (FRS-r 44)	1 month	Dysarthria / -	-
S16/M	66	Spinal-onset ALS (FRS-r 41)	2 months	Muscular weakness (left leg) and slight dysarthria / -	Crutch
S17/M	78	Spinal-onset ALS (FRS-r 40)	5 months	Hands and arms weakness and slight dysarthria / -	-

were seated at their own wheelchairs (see a photo taken during one experimental session in Fig. 6.2). The remaining disabled and able-bodied participants were seated on a standard chair. The experimental protocol and EEG recording is summarized in Table 6.3.

During the calibration phase, the participants attended the letters of the word 'INTERFACE' (9 characters) which were successively provided at the top of the screen in

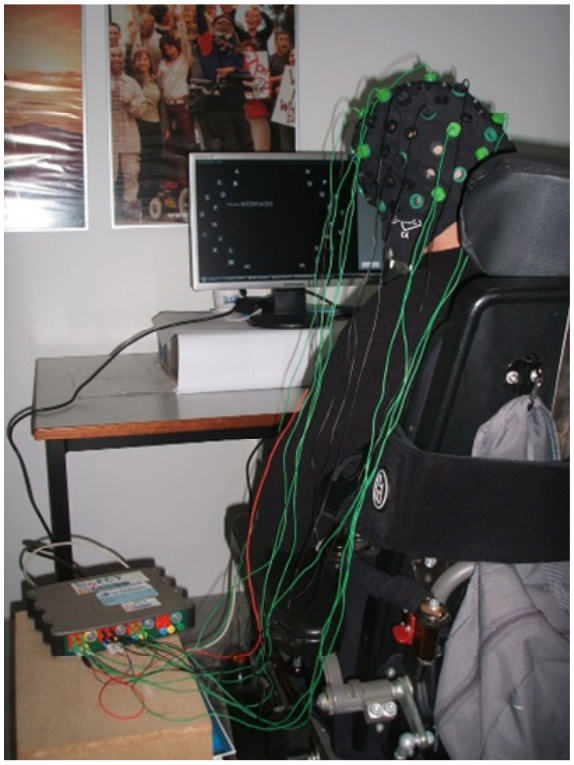


Figure 6.2: Photo taken at APCC during an experimental session.

Table 6.3: *Case-study II*: summary of experimental protocol

Recording	
Channels:	12 Ag/Cl electrodes Fz, Cz, C3, C4, CPz, Pz, P3, P4, PO7, PO8, POz and Oz (10-20 extended positions) Reference: left or right ear lobe; Ground: AFz
Frequency sampling:	256 Hz
Preprocessing filters:	0.1-30 Hz bandpass filter and 50 Hz notch filter
Impedance:	$< 10K\Omega$
Paradigm protocol	
Paradigm:	6×6 matrix RC speller
RC parameters:	SOA=200 ms; ISI=100 ms, ITI=3.5 s
Screen:	15" at about 60-70 cm
—	
Paradigm:	LSC speller
LSC parameters:	virtual ISI=75 ms, ITI=3.5 s
Screen:	15" at about 60-70 cm
Participants	10 able-bodied, 5 CP, 1 DMD, 1 SCI and 7 ALS

the RC speller, and at the center of the screen in the LSC speller. For each calibration letter, each symbol was flashed 10 times. The recording session of the calibration was less

than 5 minutes for both paradigms. The calibration dataset of each participant contains 180 target epochs and 900 non-target epochs for the RC speller, and 90 target epochs and 2430 non-target epochs for the LSC speller. These datasets were used to obtain the classification models. The online sessions occurred under the same conditions as the calibration phase. There was no difference in the procedure between the able-bodied participants and the impaired participants, except for subject S14. This participant evidenced a low visual acuity and it was necessary to point out the location of each letter. It was asked to the able-bodied participants to spell the sentence 'THE-QUICK-BROWN-FOX', and it was asked to the motor impaired participants, not familiarized with the English language, to write the portuguese sentence 'ESTOU-A-ESCREVER'. The sentences were written at once without interruptions. In case of error, participants could opt to correct the character using the 'del' symbol. The duration of each experiment depended on the number of sessions, which in turn, varied according to the participant's performance.

6.4.2.1 Classification

The offline and online classification was performed using the statistical spatial filter C-FMS combined with the qNB classifier presented in chapter 5. The spatial filter and classification models were obtained for each participant from the calibration data. In the RC speller the whole calibration data set was used (180 target epochs and 900 non-target epochs), while in the LSC speller it was used the 90 target epochs and 840 of the 2430 non-target epochs (to have more balanced datasets). In RC, the online selection of the symbol resulted from the combination of the row and column with highest scores using (5.36). In LSC, the detected symbol was the one associated with the event with the highest score, using an equation adapted from (5.36) for the 28 events.

6.5 Results

6.5.1 Online accuracy and bit rate

Online results consisting on accuracy, number of event repetitions (N_{rep}), symbols per minute (SPM) and ITR in bits per minute (bpm) for each paradigm are shown in Tables 6.4, 6.5 and 6.6. The last column of Tables 6.4 and 6.5 and the second last column of Table 6.6 indicate the order by which the two paradigms were tested. Each participant performed a series of sessions that depended on his performance. If the participant had a good performance, the number of repetitions was decreased and the whole sentence was

spelled again. If the performance was weak, then the number of repetitions was increased. The notes (a), (b), (c) and (d) in Tables 6.4, 6.5 and 6.6, describe particular occurrences during online sessions, namely (a) test not performed or skipped; (b) test aborted due to excessive number of errors; (c) test not performed because offline classification was unable to detect P300 with an accuracy above 80%; (d) test aborted due to user visual discomfort. The performance metrics, ITR, PBR and online P_{ac} were computed according to section 3.2.1 The results presented in Tables 6.4, 6.5 and 6.6 consider the ITI time, however for comparison with other research groups, the averages were also computed omitting the ITI (referred in tables as $average^{(2)}$). For a proper comparison of the two paradigms, the averages were computed by choosing, for each participant, the number of repetitions (N_{rep}) that provides the highest PBR[1], i.e., considering that all mistakes would be corrected.

Five participants were unable to effectively control RC, namely S05, S15, S19, S21 and S22, and six participants were unable to effectively control LSC, namely S14, S15, S19, S20, S21 and S22. For group I (excluding participant S05), the ITR results showed a 5.52 bpm improvement of LSC over RC (statistical paired *t-test*, $p < 0.001$). For group II, the ITR of LSC was 3.74 higher than for RC, but the statistical test only approached statistical significance ($p = 0.074$, t-test performed excluding participants S14 and S15). For group III, the ITR of RC was 1.48 higher than for LSC, but the difference was not statistically significant (t-test performed excluding participant S19, S20 and S21). Comparing the three groups, LSC was on average better for group I and group II, and RC was better for group III. The ITR average taking together the three groups was 26.11 bpm (89.90% accuracy, 4.47 N_{rep} and 6.58 SPM) for LSC, and 21.91 bpm (88.36% accuracy, 4.82 N_{rep} and 5.26 SPM) for RC. The ITR difference, 4.20 bpm, was statistically significant, $t(16) = 3.62$, $p < 0.002$ using a paired *t-test*, excluding participants S05, S14, S15, S19, S20, S21 and S22.

For group III, it was asked to participants to write a sentence with the same number of characters of the one tested on the BCI experiments, using their usual non-EEG interfaces. The last column of Table 6.6 shows the achieved SPM. These experiments were performed on a different day of the BCI experiments. Participants used different systems according to their main disability profile, namely, a head-tracker (HT), a scanning-switch (SS), a head-pointing-device (HPD), a mouse controlled by the foot (FM) and a keyboard-grid (KG). Participants S19, S21, S22 who were unable to control the BCI are the ones who reached

[1]It should be noted that the PBRs are not shown in Tables 6.4, 6.5 and 6.6, but they were computed to select, for each participant, the N_{rep} taken for average.

the highest SPMs. These communication rates are much higher than the ones achieved in current BCI systems. Participants S18, S20 and S24 achieved an SPM approximately two to three times the SPM achieved with BCI. Participant S23, who is currently unable to control any interface, achieved a BCI performance of 4.28 SPM with 87.5% accuracy.

Table 6.4: Online results of able-bodied participants (group I). See text for notes (a), (b), (d).

		RC speller			LSC speller			
Subject	N_{rep}	SPM	P_{ac} (%)	ITR (bpm)	SPM	P_{ac} (%)	ITR (bpm)	order (1st/2nd)
S01	5	3.87	89.47	16.04	4.28	90.00	16,55	RC/LSC
	4	4.58	85.00	17.36	5.04	90.47	19.66	
	3	(b)	-	-	6.12	81.81	19.95	
S02	5	3.87	86.60	18.63	(a)	-	-	LSC/RC
	4	(b)	-	-	5.04	100.0	24.23	
	3	(a)	-	-	6.12	90.47	23.88	
	2	(a)	-	-	7.79	81.81	25.39	
S03	7	2.95	84.21	11.02	3.29	90.47	12.85	LSC/RC
S04	5	3.87	82.60	13.97	4.28	90.00	16.55	LSC/RC
S05	5	(d)	-	-	4.28	89.47	16.37	LSC/RC
S06	5	3.87	100.0	20.01	4.28	94.737	18.25	LSC/RC
	4	4.58	84.21	17.08	4.28	95.00	21.59	
S07	5	3.87	95.00	17.91	4.28	89.47	16.37	RC/LSC
	4	4.58	90.47	19.36	(a)	-	-	
	3	5.60	79.16	18.85	6.12	85.71	21.65	
S08	5	3.87	85.00	14.17	4.28	90.47	16.71	RC/LSC
	4	4.58	70.37	12.70	5.04	86.36	18.07	
	3	(a)	-	-	6.12	90.00	23.65	
S09	5	3.87	100.0	20.01	4.28	90.47	16.71	RC/LSC
	4	4.58	86.36	17.84	(a)	-	-	
	3	5.60	90.47	23.70	6.12	95.00	26.22	
	2	7.22	95.00	33.44	7.79	100.0	37.45	
S10	5	3.87	100.0	20.01	4.28	100.0	20.60	RC/LSC
	4	4.58	90.47	19.36	5.04	100.0	24.23	
	3	5.60	95.00	25.94	6.12	95.00	26.22	
	2	(b)	-	-	7.79	85.71	27.55	
Average(1)	4.44/3.90	4.49	89.32	18.94	5.48	91.68	22.16	
Average(2)	4.44/3.90	5.64	89.32	23.97	7.29	91.68	29.49	

(1) Average of the elements in the table (includes ITI).
(2) Average if ITI would be excluded.

Table 6.5: Online results of ALS participants (group II). See text for notes (a) and (b).

		RC speller			LSC speller			
Subject	N_{rep}	SPM	P_{ac} (%)	ITR (bpm)	SPM	P_{ac} (%)	ITR (bpm)	order (1st/2nd)
S11	5	3.87	100.0	20.01	4.28	87.50	15.72	RC/LSC
	4	4.58	81.25	16.08	5.04	87.50	18.50	
	3	(b)	-	-	6.12	87.50	22.46	
S12	5	3.87	93.75	17.46	4.28	93.75	17.88	RC/LSC
	4	4.58	87.50	18.25	5.04	100.0	24.23	
	3	5.60	87.50	22.34	6.12	81.25	19.71	
S13	7	2.95	93.75	13.33	3.29	93.75	13.75	LSC/RC
	5	3.87	68.75	10.33	(b)	-	-	
S14	7	2.95	100.0	15.28	3.29	43.75	3.77	LSC/RC
	6	3.35	68.75	8.95	(a)	-		
S15	≥ 7	(b)	-	-	(b)	-	-	LSC/RC
S16	4	4.58	87.50	18.25	5.04	100.0	24.23	LSC/RC
	3	5.60	87.50	22.34	6.12	93.75	25.54	
	2	(b)	-	-	7.79	73.33	21.06	
S17	5	3.87	81.25	13.59	4.28	81.25	13.79	RC/LSC
Average(1)	5.0/4.40	4.14	91.67	17.82	4.97	91.25	19.95	
Average(2)	5.0/4.40	5.10	91.67	21.83	6.37	91.25	25.57	

(1) Average of the elements in the table (includes ITI).
(2) Average if ITI would be excluded.

6.5.2 Questionnaire

Users were asked to answer a questionnaire to assess the degree of user satisfaction and to identify limiting factors for both paradigms. Participants with high dysarthria in the ALS group answered the questionnaire with the assistance of relatives. Participants of group III answered the questionnaire with the assistance of occupational therapists. The results of the questionnaire are in Table 6.7. The answers to each specific question/item are shown as the percentage over the sample of the participants. The results show that most of the participants liked more the LSC speller. This subjective answer has no straightforward relation with user performance. There were cases in which the user liked more one of the interfaces and yet had a worse performance. Also, 50% of the participants reported that RC speller caused higher eyestrain and discomfort than the LSC speller, causing for instance 'weeping' and 'visual after-effect'. Participants reported that the ITI should be increased (50% for RC and 31.8% for LSC), and that some errors might have occurred because of this limited time. The difficulty to find the letters according to the layout was higher for RC. Participants were asked to mention other effects that affected their performance. Three items (mental counting, distractors and double flash) were reported

Table 6.6: Online results of participants CP, DMD and SCI (group III) and performance with non-EEG interfaces. The non-EEG interfaces SS, HT, HPD, FM, KG, refer respectively to scanning-switch, head-tracker, head-pointing-device, foot-mouse and keyboard-grid. See text for notes (a), (b) and (c).

		RC speller			LSC speller				Non-EEG interface
Subject	N_{rep}	SPM	P_{ac} (%)	ITR (bpm)	SPM	P_{ac} (%)	ITR (bpm)	order (1st/2nd)	SPM
S18	8	(a)	-	-	2.95	94.75	12.33	LSC/RC	SS: 6.8
	7	(a)	-	-	3.29	81.25	10.61		
	6	3.35	100.0	17.32	3.72	68.75	9.0		
	5	3.87	88.88	15.85	(a)	-	-		
S19	(c)	-	-	-	-	-	-	-	FM: 43.2
S20	6	3.35	93.37	15.12	3.72	50.00	5.32	RC/LSC	HT: 9.6
S21	≥ 9	(c)	-	-	(b)	-	-	RC/LSC	HPD: 51.4
S22	(c)	-	-	-	-	-	-	-	KG: 31.7
S23	7	(a)	-	-	3.29	93.75	13.75	LSC/RC	HT: 0.0
	6	3.35	68.75	8.95	3.72	87.50	13.67		
	5	3.87	75.00	11.90	4.28	87.50	15.72		
S24	8	2.64	73.33	7.83	2.95	56.25	5.13	RC/LSC	HT: 7.2
	7	2.95	37.50	2.98	3.29	62.50	6.82		
Average(1)	6.25/6.67	3.30	85.43	13.04	3.51	81.58	11.62		
Average(2)	6.25/6.67	3.84	85.43	15.18	4.13	81.58	13.70		

(1) Average of the elements in the table (includes ITI).
(2) Average if ITI would be excluded.

by their own initiative. For the RC speller, 55% of the participants reported that the frequent mental counting of the target events was tedious and somehow made them lose focus. This perception experienced by the participants may be explained by the fact that in the RC speller, users have to count twice (one for target row and one for target column) for each round of flashes, while in LSC speller, users have to count only once for each round of flashes (target symbol). The double-flash effect was also reported only for the RC speller by 13.6% of the participants. The adjacency distractors were pointed out for both paradigms, but with higher incidence for the LSC speller. Participants were asked to propose modifications and improvements to the paradigms. For RC, it was suggested to soften the colors. For LSC, it was suggested to increase the distance between symbols, to increase the symbols size, and to change the style of the letter "I" to improve its perception.

Table 6.7: Results of the questionnaire.

Question/item	RC	LSC	Similar	None
Paradigm most appreciated	13.6%	72.7%	13.6%	n.a.
Paradigm that caused more eyestrain/visual discomfort	59.0%	0.0%	n.a.	41.0%
ITI is not long enough	50.0 %	31.8%	n.a.	50.0%
Symbols are difficult to find	50.0%	18.2 %	n.a.	27.3%
Undesired effects reported by users				
High mental counting	55%	0%	n.a.	n.a.
Double-flash	13.6%	0%	n.a	n.a.
Distractors	4.5%	13.6%	n.a.	n.a.

n.a.: not applicable

6.5.3 Adjacency errors

To analyse the influence of the adjacency distractors, the number of errors was computed according to spatial location of the respective events. This information was collected only from the disabled participants. For the RC speller, 90% of the errors were on the same row or on the same column (row: 38% and column 52%). Moreover, 54% of the errors were in one of the 8 adjacent symbols. In the LSC speller, 32% of the errors were in one of the 4 adjacent symbols. Also, 47% of the errors were on the same side and 53% in the opposite side. These results may indicate that most of the RC errors are due to adjacency distractors. In LSC, there is also a high percentage of adjacency errors, yet the percentage is considerably lower than in RC, which may indicate that the errors are mostly due to P300 variability, or due to remote distractors.

6.6 Offline analysis

An offline analisys was carried out to understand and sustain the online results, but also to assess if our initial hypotheses were correct. The analysis was based on the datasets collected during the calibration phases.

6.6.1 Waveform morphology

The analysis was focused on channels Pz and PO7, two of the most discriminative channels [Kaper 2004, Krusienski 2008, Townsend 2010], and focused on the N200 and P300 ERP components, since they are the two main discriminative components. Fig. 6.3 shows the average of target epochs for each participant with both paradigms. For a better visualization, the epochs were low-pass filtered in the time domain at $f_c = 7$ Hz (it should

be noted that the filter causes a temporal delay of ≈ 50 ms). The analysis hereinafter, however, uses raw EEG data, i.e., without filtering. The components of the ERP elicited in the RC speller, namely the evoked N200 and P300 components are sometimes difficult to discriminate. This might happen because the signal is strongly affected by a steady state visual evoked potential (SSVEP) that results from the 5 Hz stimulus flash, and also because target epochs temporally overlap due to the small TTI. In the LSC speller, the N200 and P300 components are usually clearly identified. Peak amplitude and peak latency were determined by selecting the largest positive or negative peak within the range 275-600 ms (P300) and 175-400 ms (N200). From the individual analysis, we concluded that participant S19 did not elicit any visible ERP component for both paradigms. Participant S22 exhibits components with very high amplitudes which reveal the presence of artifacts. These two participants were excluded from the analysis. Participants S15 and S17 exhibit a strong positive peak around 200 ms, but the P300 component is hardly noticeable. Nevertheless, they were not excluded from the analysis.

A summary of the average amplitudes and latencies of P300 and N200 peaks, as well as the respective statistical tests (*t-test*) is provided in Table 6.8. The P300 amplitude is larger for LSC than for RC and the latency is shorter for LSC than for RC. The statistical test fails only for the amplitude difference in channel PO7. For the N200 component, the amplitude is larger for RC, but the statistical test only approaches significance for channel Pz. The latency difference is not statistically significant.

The grand averages were computed for group I, group II, and participants with CP (only S18, S20 and S21), as shown in Fig. 6.4. Qualitatively, it can be seen that the ERP waveform elicited by the LSC paradigm shows a clear classical morphology, particularly in able-bodied and CP participants. In RC, the waveforms morphology exhibit a more pronounced oscillatory component. The amplitudes and latencies of the P300 component (raw EEG data) of the able-bodied were statistically compared with those obtained with the ALS and CP groups, using a two-sample *t-test*. It should be noted that the sample size of the CP group is very small and therefore the statistics should be taken with care. The P300 amplitude was similar across groups and the latencies are shorter for the able-bodied participants, approaching statistical significance in RC speller (both channels), and being statistically significant between able-bodied and ALS in the LSC speller paradigm (PO7 channel).

The question of whether the P300 systems are gaze dependent has been a topic of discussion in the last few years [Kaper 2004, Krusienski 2008]. It is today assumed that most of the P300-based BCI systems resort to overt attention, and that they use not only

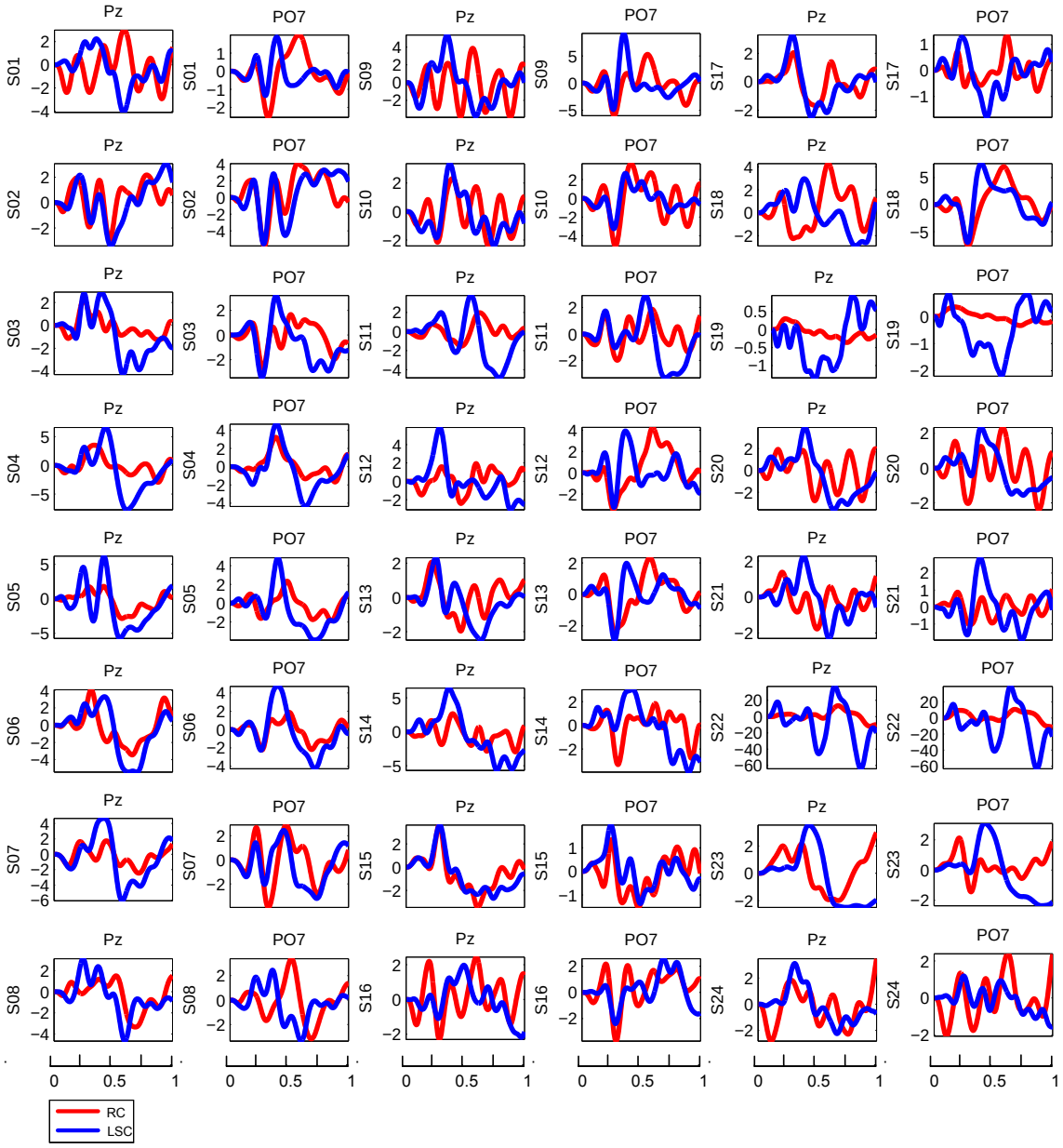


Figure 6.3: Averaged target epochs elicited by RC paradigm (red color) and LSC paradigm (blue color) at channels Pz and PO7 for each of the 24 participants (amplitudes units are μV). All waveforms were low-pass filtered ($fc = 7$ Hz) before plotting.

the P300 component but also other specific components such as the N100, P200 and N200 components that are modulated by attention tasks and influenced by gazing [Treder 2010, Frenzel 2010]. Our own observations, show that these components, in particular the N200 component, provide valuable features and play an important role to increase the classification accuracy. For this reason, it has been suggested that P300-based BCIs

Table 6.8: Mean peak amplitudes and latencies of P300 and N200, LSC and RC differences, and statistical analysis.

Component (channel)	Property	Diff. between LSC and RC (stat. test)
P300 (Pz)	Amplitude LSC: 4.47 μV RC: 2.88 μV	dif = $+1.59\mu V$ ($t(21) = 3.1$, $p < 0.004$)
	Latency LSC: 341 ms RC: 415 ms	dif = -74 ms ($t(21) = 3$, $p < 0.006$)
P300 (PO7)	Amplitude LSC: 4.28 μV RC: 3.61 μV	dif = $+0.67\mu V$ (no stat. sign.)
	Latency LSC: 354 ms RC: 476 ms	dif = -122 ms ($t(21) = 6.2$, $p < 0.001$)
N200 (Pz)	Amplitude LSC: -1.71 μV RC: -2.35 μV	dif = $-0.64\mu V$ ($t(21) = 1.9$, $p = 0.07$)
	Latency LSC: 295 ms RC: 320 ms	dif = -25 ms (no stat. sign.)
N200 (PO7)	Amplitude LSC: -3.67 μV RC: -3.71 μV	dif = $-0.04\mu V$ (no stat. sign.)
	Latency LSC: 239 ms RC: 230 ms	dif = $+9$ ms (no stat. sign.)

should be renamed 'ERP-based' BCIs [Treder 2010].

6.6.2 SNR contribution to ERP variability analysis

Several studies quoted in section 6.2.2 show non-consensual results suggesting the absence of a straightforward relation between P300 amplitude and classification accuracy. SNR measurement of target ERPs can add valuable information to their characterization in addition to the waveform morphology, since it assesses the variability of inter-trial ERPs. Although classification accuracy depends not only on the ERP variability, but also on the separation between target and non-target ERPs, our observations in chapter 5 indicated a direct relationship between SNR and classification performance. To infer this relationship and to assess the inter-trial variability of ERPs, the SNR was computed for both paradigms using (3.14).

Fig. 6.5 shows the SNR averages (in dB) separately for able-bodied, ALS and CP participants, taking $K = 1 \cdots 7$ epochs for average. As expected, and in line with the

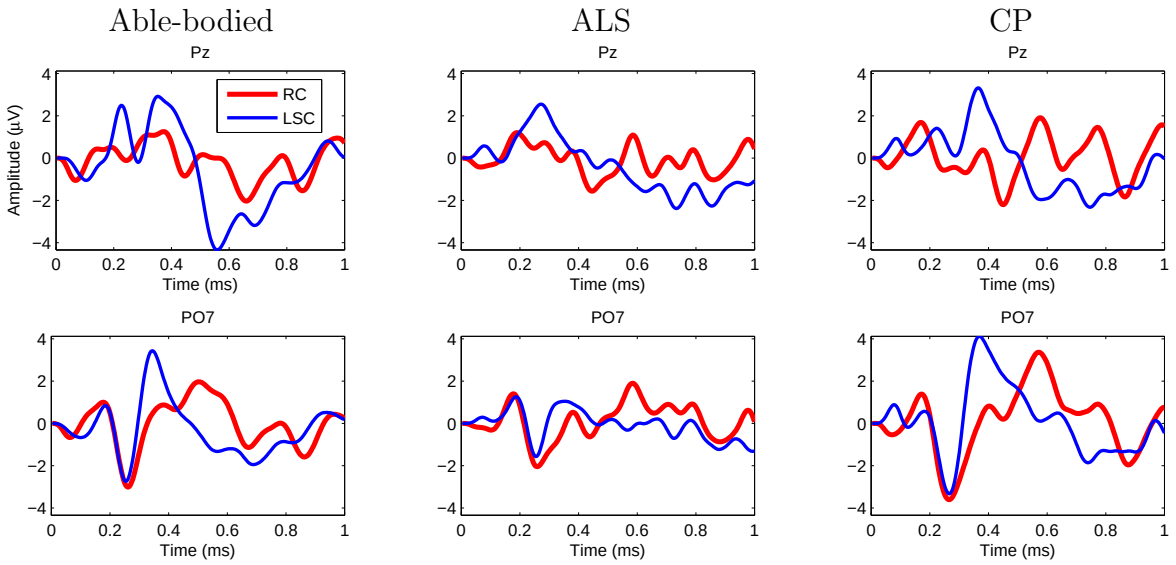


Figure 6.4: Grand averages computed over able-bodied group, over ALS group and over CP group (only S18, S20 and S21). All waveforms were low-pass filtered ($fc = 7$ Hz) before plotting.

online results, the able-bodied group has the highest SNR for both paradigms, and the CP group has the lowest SNR, despite the fact that the CP participants exhibited a higher P300 amplitude. Comparing RC and LSC paradigms in the able-bodied group, the SNR of LSC is always at least 3.7 dB higher than RC, for $K = 1 \cdots 7$, p <0.006. The ALS group also shows a difference > 3.7 dB for $K = 1 \cdots 7$, but it only approaches statistical significance, $p < 0.08$. In the CP group (S18, S19 and S20), the difference is > 2.7 dB, but it is not statistically significant.

6.6.3 Searching for new discriminative features in LSC

The LSC paradigm was designed to overcome some limitations of the RC paradigm, as well as to investigate the emergence of new neurophysiologic features. For this purpose, the existence of discriminative features in the frequency domain was investigated, and also whether new neurophysiologic features could arise from the left/right layout and event presentation strategy of LSC.

6.6.3.1 Frequency analysis

As the user focuses on only one side of the screen, it was initially hypothesized that the stimuli on the opposite side would have a small visual effect, and thus the user would

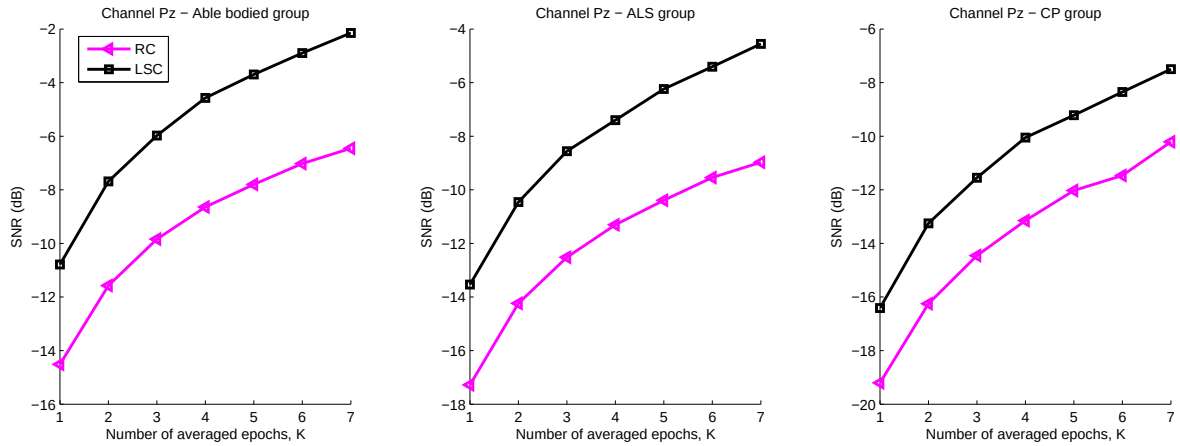


Figure 6.5: SNR estimated for able-bodied group, ALS group, and CP group (only S18, S20 and S21). The SNR was computed for $K = 1 \cdots 7$ averaged epochs.

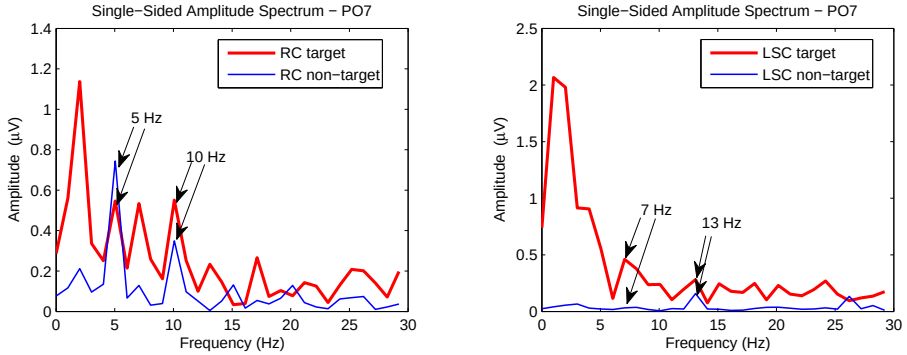


Figure 6.6: Frequency analysis illustrating the inter-stimulus flashing effect on RC and LSC. The plots correspond to the averages of 180 target FFTs and 900 non-target FFTs for RC, and 90 target FFTs and 840 non-target FFTs for LSC. The averages were computed from datasets chosen from a representative participant.

virtually see a target probability of 1/14 and only half of the stimuli, having a perception of 6.5 Hz, i.e., half of the flashing frequency. Fig. 6.6 corroborates this assumption. Target epochs reflect a frequency at 6.5 Hz (7 Hz in the graph due to FFT resolution), and only a small peak at 13 Hz. Non-target epochs reflect only the 13 Hz flash. By contrast, the RC paradigm shows that the frequency of the flashes strongly affects both target and non-target epochs of the RC speller. The LSC graph in Fig. 6.6 suggests that LSC may have discriminative features related to stimuli frequency that can be used to improve the classification. To infer the influence of frequency, we computed the offline classification accuracy using the SSFCB filter (5.27) that includes a frequency shaping (see chapter 5), and we compared it with C-FMS. Fig. 6.7 shows the averaged classification

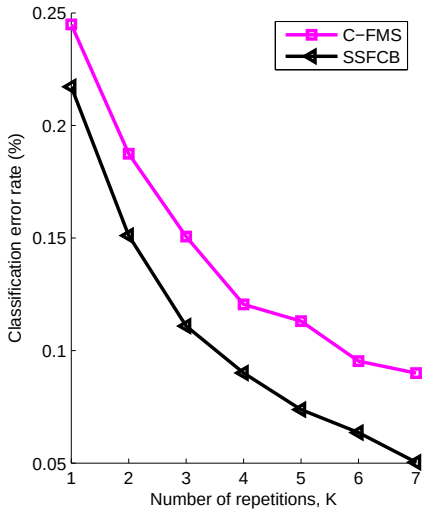


Figure 6.7: Offline classification results averaged over all participants except S19 and S21. Results obtained with C-FMS filter (which was applied online) and with SSFCB filter that incorporates a spectral shaping.

error rate over all participants except S19 and S21 obtained with both methods (using unseen data). Results show that the SSFCB achieves a lower error rate (statistical t-test p <0.05 for $K = 1 \cdots 7$), suggesting that LSC speller can benefit from the spectral shaping. The filter $\tilde{\mathbf{h}}$ (5.25) was manually adjusted setting to zero the weights associated with non-discriminative frequency bins.

6.6.3.2 Laterality analysis

Visual spatial attention modulates early ERP components such as the N100, P200 and N200 with lateralization effects [Luck 1994, Makeig 1999]. However, experiments are usually performed based on covert attention, while in our case an overt attention paradigm was used. By computing the ERP waveforms elicited by left and right events, we investigated whether or not a laterality difference exists. Fig. 6.8 (top) displays the ERP waveforms of epochs recorded at PO7 and PO8 elicited by left and right target events. The waveforms are the grand averages computed over the group of able-bodied participants. Components N100, P200 and N200 of the ERP exhibit different waveforms for left and right events. The difference is more easily detected by computing the difference between PO8 and PO7 as shown in Fig. 6.8 (middle). For example, the component N200 is larger over PO7 for both left and right events, but the PO8-PO7 difference is larger for right events. The results suggest that the classification algorithms could benefit from

two individual models, one for left events and the other for right events. Computing the PO8-PO7 difference for non-target epochs, a phase reversal on the SSVEP is evidenced in Fig. 6.8 (bottom). This phase reversal, which is produced by the left/right event strategy of the LSC paradigm, indicates that it is possible to detect the side of the event only by computing the phase between EEG signals from left/right symmetric locations.

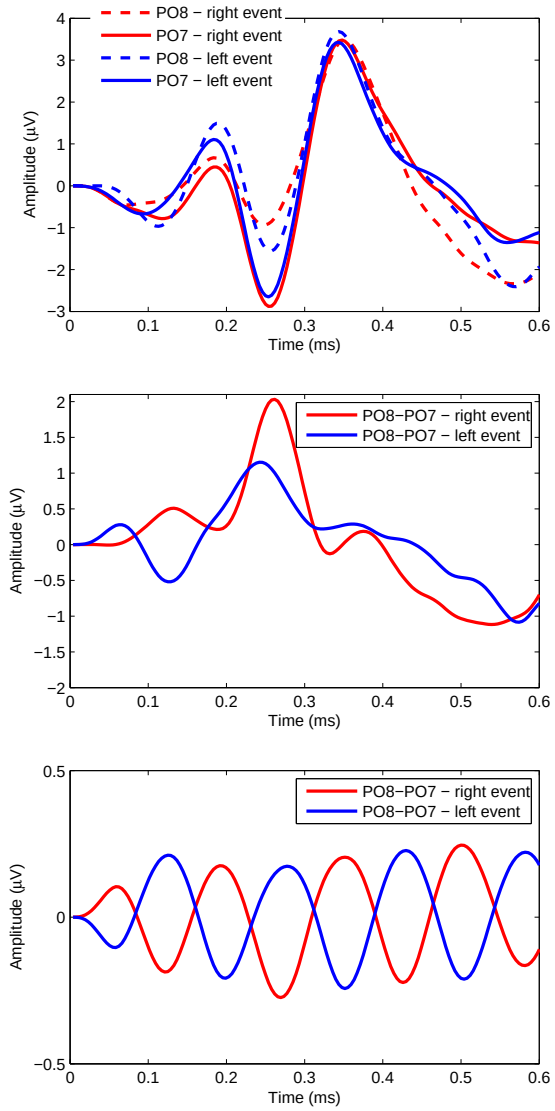


Figure 6.8: Grand averages computed over able-bodied participants. Top: target epochs recorded at PO7 and PO8 for left and right events; Middle: difference between target epochs recorded at PO8 and PO7 for left and right events; Bottom: difference between non-target epochs recorded at PO8 and PO7 for left and right events. The curves show a phase reversal effect.

6.7 Discussion

The initial assumptions regarding the LSC paradigm, indicating that this approach could be an effective alternative to the RC paradigm, were assessed by online and offline analysis, and deserve now some further discussion.

6.7.1 Online performance and state-of-the-art comparison

The average results (including ITI) of the able-bodied participants were 18.94 bpm, 89.32% accuracy, 4.49 SPM and 4.44 repetitions for RC, and 22.16 bpm, 91.68% accuracy, 5.48 SPM and 3.90 repetitions for LSC. These results compare very favourably with those reported in the two studies with the highest performances. Namely, [Townsend 2010] reports a 19.85 bpm, 77.34% accuracy, 4.68 SPM and 4.5 repetitions for RC, and 23.17 bpm, 91.52%, 4.36 SPM and 3.61 repetitions for the checkerboard paradigm, and [Lenhardt 2008] reports a 3.91 SPM and 83.33% accuracy for RC. The online results achieved by motor impaired participants are also good in comparison with the results reported in other studies [Townsend 2010, Hoffmann 2008b, Sellers 2006b, Piccione 2006, Donchin 2000], however results cannot be directly compared because the levels of disability of the participants greatly differ.

Comparing LSC and RC, the overall averaged ITR (excluding only the subjects unable to participate online in one or both paradigms) was 4.20 bpm higher for LSC. The relevant improvement of LSC over RC, is a major achievement and shows that LSC is an effective alternative to RC. Nonetheless, while the ALS participants performed better with LSC showing only slightly worse results than those achieved by able-bodied participants, the participants of group III, in particular the CP participants, performed better with RC. The overall results obtained with CP participants are also relevant since there are very few studies examining the use of P300-based BCIs by individuals with CP [Hoffmann 2008b]. From the group of five CP participants, only two effectively controlled the BCI, namely S18 and S20. The overall results obtained with CP participants may suggest that these individuals may have more difficulties to control a BCI than the other disabled participants. In order to be more conclusive about the participants who performed worse, some participants of group III (S19, S20, S23 and S24) underwent pilot experiments with an additional paradigm presented in chapter 8 called Arrow-paradigm. The aim was to assess whether the weaker results achieved by some participants of this group were only due to neurophysiologic causes, or depended more on the complexity of RC and LSC. The Arrow-paradigm has a simpler presentation than RC and LSC, consisting of only 11 large

and well separated symbols that flash individually. Using this paradigm, participant S19 evoked a P300, although it was still weak and insufficient for online operation. Participants S20 and S24 significantly improved the online performance and S23 kept the online performance. These results suggest that it is worth to test different visual paradigms to increase the BCI performance.

6.7.2 Contributions of the LSC paradigm

The event strategy chosen for LSC allowed to significantly decrease the flashing time and to eliminate the ISI (while still keeping a virtual ISI effect). Thereby, the reduction of the time associated with the events allowed to achieve sub-trial times similar to those typically obtained with RC, leading to competitive ITRs. By reducing the surrounding symbols in LSC, it was possible to control the local distractors and thereby to reduce the adjacency errors. This percentage was 32%, while in RC it was 54%. The percentage of remote errors, in the opposite side of the target event, was 53%. We do not have enough information to conclude whether the error percentage is due to ERP variability or due to remote distractors, but according to the questionnaire, participants reported that distractors affected their performances. Other event strategies, for example, extending the left/right event strategy to the four quadrants of the screen, can potentially be explored in the future in combination with further manipulations of TTI, ISI and double-flash.

The questionnaire allowed us to assess some subjective issues concerning RC and LSC. The majority of participants, 72.7%, expressed a preference for LSC. Additionally, 59% reported that RC caused a higher eyestrain and visual discomfort, while in LSC none of the participants reported any visual discomfort. This is an important issue in BCI and should not be overlooked because it may limit the time period of BCI usage.

6.7.3 Neurophysiologic features

The grand-averages in Fig. 6.4 show that the ERP waveforms for both RC and LSC differ between Pz and PO7. The P300 amplitude in channel Pz is larger in LSC than in RC. The peak occurs also earlier in LSC than in RC, for channels Pz and PO7. These results corroborate our initial expectations that LSC, with a lower target probability, would elicit a larger P300 amplitude than RC. As concerns N200, the amplitude in Pz was larger in RC than in LSC, suggesting that the N200 component may have contributed more to the classification accuracy in RC than in LSC.

When searching for features induced by the lateral layout and event strategy of LSC,

the grand averages suggested a discrimination between left vs. right ERP targets. This discrimination can be exploited, by using left and right independent models in the classifiers. These asymmetries are a direct reflection of the hemispheric dominance of the right hemisphere in visuospatial attention [Silva 2008, Silva 2010]. Other asymmetries (e.g. up/down) in the paradigm design can potentially be explored in the future.

A major achievement of LSC paradigm was the SSVEP phase reversal effect between left and right non-target events. This effect can enable the detection of the side of the target event by measuring the phase at symmetric electrode locations. The inclusion of phase information adds a new dimension to the P300-based BCI that can increase the classification accuracy. The phase information can be embedded in current classification algorithms or combined in a different classifier. Detecting the side of the incoming event can be used to discard the events from the opposite side, or to increase the estimated likelihood of the target event.

6.7.4 Limitative factors of BCI

BCI is still limited by various factors. It usually requires more attention and may cause a higher fatigue than other standard interfaces. Attention tasks may be affected when users experience pain symptoms. For example, some of the participants of group III had to interrupt the sessions to alleviate the pain, by stretching the legs, arms and shoulders, or by changing the position of the head, which may have negatively influenced their performance. Curiously, participants of group III who were unable to control the BCI, were those who had better results with standard interfaces. Motivation may eventually play an important role on the achieved performances. Some participants, who are accustomed to use efficiently the same interface for more than a decade, were not highly motivated. Moreover, it was subjectively observed that the level of confidence and the fear to fail could influence the performance. Please see the study in [Nijboer 2010] where some of these issues are discussed.

6.8 Conclusion

This study presented a P300-based paradigm, the LSC speller, that showed an increase in performance over the standard RC speller. LSC paradigm explored the event strategy and the hemispheric asymmetries to introduce new neurophysiologic dimensions that can be used in the future to improve bit rates and classification accuracy. The effective transfer

rates achieved by motor disabled participants for both LSC and RC spellers, are a step forward to their potential application as assistive devices for communication.

Chapter 7

GIBS: Gaze Independent Block Speller

Contents

"Everything having to do with human training and education has to be reexamined in light of neuroplasticity."

Norman Doidge, in "The Brain That Changes Itself".

This chapter introduces a novel two-level speller paradigm called gaze independent block speller (GIBS) and compares it with the RC speller. Online experiments show that the users can effectively control GIBS without moving the eyes (covert attention), while this task is impossible with the RC speller. Additionally, in overt experiments (gazing the target), GIBS attains performances similar to RC, despite being a two-level speller. This chapter extends also the practical bit rate (PBR) metric for the case of two-level spellers.

7.1 Introduction

A P300-based BCI is considered an independent BCI because it can be controlled without resorting to any kind of movement. While this is actually true, most of the P300 BCIs cannot however be effectively controlled without gazing the target symbol (overt attention). There is a fall off in spatial resolution of the human visual system away from the point of fixation which makes difficult the detection of target events outside the central vision. At 10 degrees from the center of the macula (fovea), eyes are only capable of seeing about 20% in terms of visual acuity (see Fig. 7.1). Gaze dependence is a limitative factor particularly relevant for individuals who lose the ability to move the eyes (e.g. advanced stages of disorders such as ALS) [Palmowski 1995]. Gaze-independent BCIs have been addressed in [Treder 2010, Liu 2011, Brunner 2010]. The authors showed that in the standard RC speller, users have to gaze the mentally selected letter to be able to control the speller. For stimuli in the periphery of the RC matrix (or simply outside the foveal vision), covert attention (without moving the eyes) is very difficult because spatial acuity decreases with increased visual eccentricity, and because the identification of symbols is hampered by the crowding effect [Treder 2010, Sampaio 2011]. To address this problem, several solutions have been proposed in [Treder 2010, Frenzel 2010, Liu 2011], trying to improve the spatial arrangement of the symbols (central layout with larger symbols) allowing a covert attention task. The authors [Treder 2010, Brunner 2010] tested the paradigms offline, and showed that the classification accuracy with covert attention was much lower than with overt attention, even with the proposed spatial layouts. Only [Liu 2011] tested the paradigm online, but successful results were achieved only for 10 repetitions, leading to a very low transfer rate.

We describe here GIBS, a gaze independent speller, which allowed effective online transfer rates when controlled covertly.

7.2 GIBS design

GIBS is a two-level speller. User has first to select a group of letters (first level) and then to select the letter within that group (second level). Screenshots of the GIBS are presented in Fig. 7.2 a) and b). The paradigm comprises all letters of the alphabet, the symbols 0 and 1, the space ('sp'), and the delete ('dl'), totalizing 30 symbols. The symbols are grouped into four blocks following an alphabetical order. The symbols 'sp' and 'dl' are repeated in every group. The symbol 'sp' is used to separate the words of the

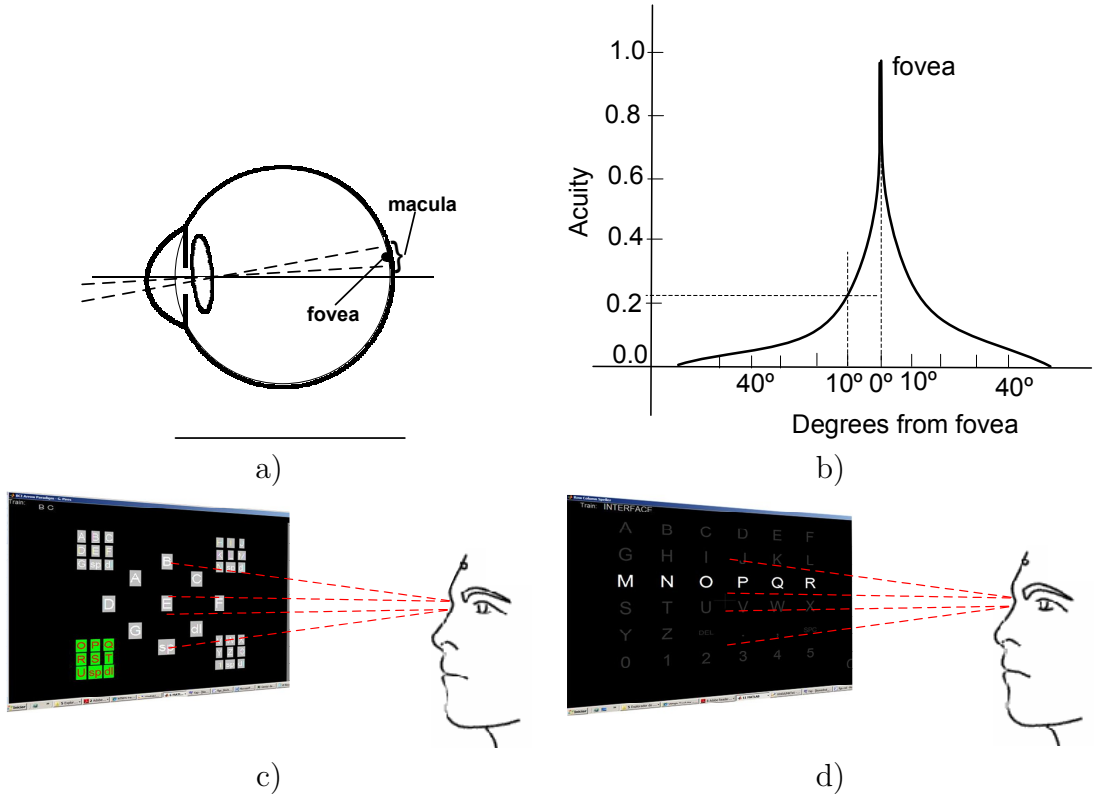
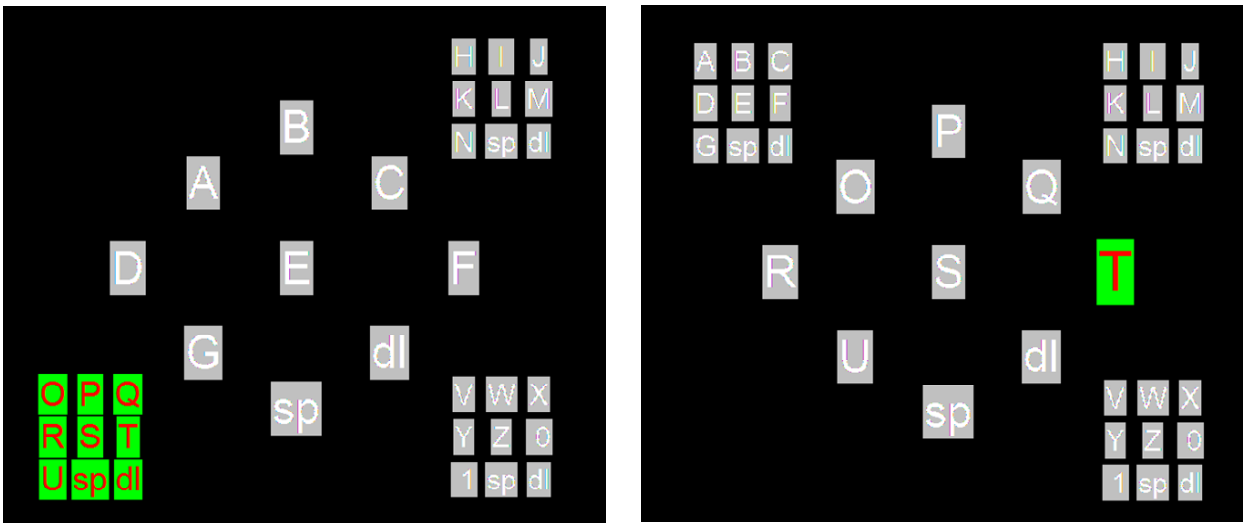


Figure 7.1: a) Simple schematic view of the fovea in human eye; b) Relative acuity vs. degrees from fovea; c) Screenshot of GIBS and field of view 5 degrees around the fovea at ≈ 60 cm of distance; d) Screenshot of RC and foveal region and field of view 5 degrees around the fovea at ≈ 60 cm of distance.

sentences. Since this symbol is frequently used, it was included in all groups to enable its selection without having to switch between blocks. The 'dl' symbol is also included in all groups with the same purpose. The paradigm layout is composed by a group of 9 symbols in the center and 3 lateral small blocks with the remaining symbols. Since there are 9 symbols at the center and 3 small blocks flashing, the number of events is 12, and thereby the target event probability is 1:12. To select a symbol, the user has first to select the small block where it belongs. When the respective block is selected, the symbols of that block move to the center and the respective small block disappears (see Fig. 7.2 a) and b)). The user can then select the attended symbol. These symbols at the center are large, well apart, and placed in the visual angle near the fovea. These symbols flash individually while the symbols within the small blocks flash as an entire block. Despite the fact that the small blocks are placed outside the central vision, this does not constitute a problem because the user has only to detect the entire block and not an individual symbol within the block. Since all blocks are always visible, in case of a missed block or misspelled



a) b)

Figure 7.2: a) Screenshot of GIBS when a block is flashed; b) Screenshot of GIBS after the lower left block has been selected.

symbol, user can immediately select the correct block or delete the symbol. The SOA was set to 150 ms, and the duration of each flash was set to 75 ms. The ITI was set to 3.5 s.

From an information transfer rate perspective, the time taken for the transitions between blocks is clearly a disadvantage over the RC speller. However, GIBS has a lower target probability, a lower occurrence of double-flash, a lower probability of overlapping, and fewer elements of distraction, reducing the crowding effect, when compared with RC. It is therefore expected that the P300 component elicited by GIBS will have a higher SNR, leading thereby to an increased classification accuracy. In sections 7.2.1 and 7.2.2, an analysis is made, showing how much the classification accuracy should be improved to achieve transfer rates similar to RC, thus compensating the transitions between blocks.

7.2.1 Block transitions rate

To estimate the time required for selecting a symbol in GIBS, we need first to estimate the average rate of transitions between blocks (N_{tr}). Taking several hundred sentences of informal dialogs and generic texts in the English language, we computed the average number of transitions that would occur using the groups of letters in Fig 7.2, i.e., using the alphabetical order. The respective groups are illustrated in Fig. 7.3a). The average number of transitions was $N_{tr} \approx 0.60$ transitions per selection. The possibility of reducing N_{tr} by using different groups of letters was also investigated. The average number of

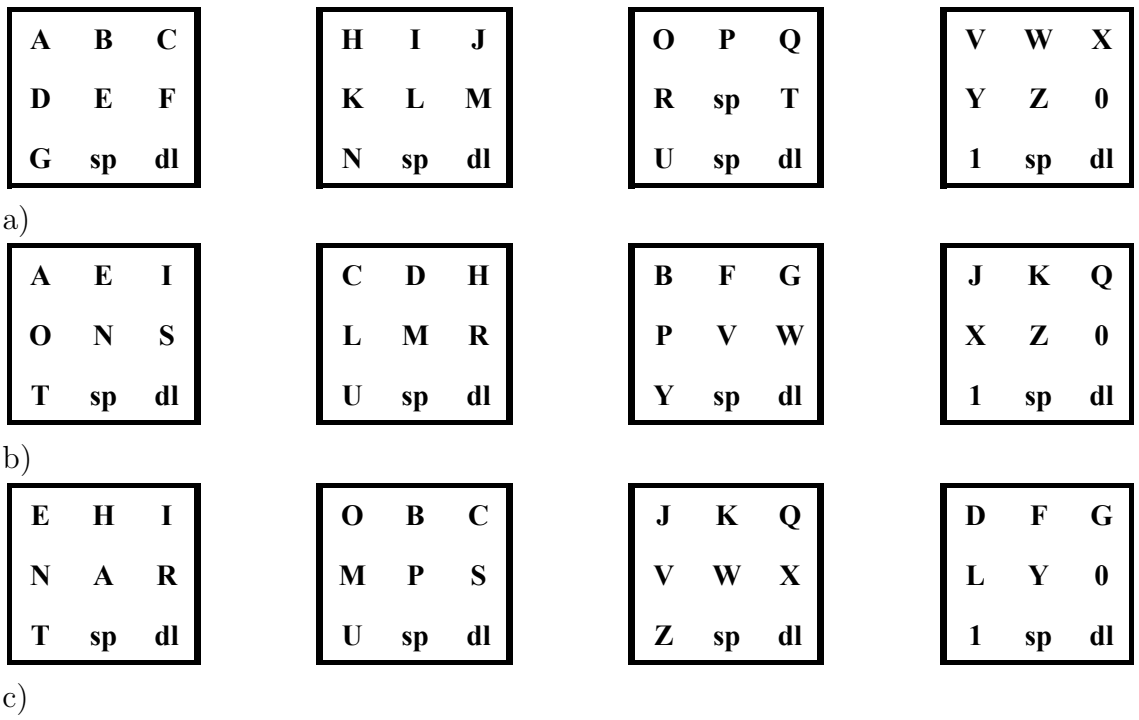


Figure 7.3: Possible groups of letters according to: a) Alphabetical order (currently used in GIBS); b) Frequency of letters (of the English language); c) Frequency of digrams (of the English language).

transitions was computed grouping the letters according to letter frequency (Fig. 7.3b)) and grouping the letters according to most frequent digrams (sequence of two letters) (Fig. 7.3c)). The achieved number of transitions were respectively 0.51 and 0.48. So, it is possible to improve the N_{tr}, however these groups are less intuitive than the ones in alphabetical order and users may experience difficulty in memorizing the letters included within each group.

7.2.2 Bit rate metrics

In section 3.2.1, we presented two transfer rate metrics, namely ITR and PBR. We will now extend the PBR metric to GIBS, which can be generalized to any two-level paradigm. Let re-equate the PBR, obtained from (3.8)

$$PBR = r_s \log_2 N_s \tag{7.1}$$

where N_s is the number of symbols and r_s is the rate of possible selections per minute, computed from

$$r_s = \frac{60}{T_d} \tag{7.2}$$

where T_d is the trial duration (time required for selecting a symbol). For the RC speller, the trial duration is given by

$$T_d = N_r(N_{rep} \times (N_{ev} \times SOA) + ITI) \tag{7.3}$$

and for GIBS, by

$$T_d = (1 + N_{tr})N_r(N_{rep} \times (N_{ev} \times SOA) + ITI), \tag{7.4}$$

where N_r is the number of tries to correctly spell a character (3.7), N_{tr} reflects the averaged number of transitions, N_{ev} is the number of events and N_{rep} is the number repetitions to make a symbol selection. Substituting (7.4) in (7.2), yields the PBR metric for the two-level speller

$$PBR = \frac{60}{(1 + N_{tr})N_r(N_{rep} \times (N_{ev} \times SOA) + ITI)} \log_2 N_s. \tag{7.5}$$

Using the above formulas (7.3) and (7.4), we can compare now the two transfer rates. For the same number of repetitions and classification accuracy, the bit rate for GIBS will be naturally lower than for the RC speller. Thus, we want to determine what should be the relative improvement in classification accuracy of GIBS to compensate the time taken for extra selections needed for transitions between the blocks. Taking (7.3) and (7.4), T_d was analyzed for two different situations. In Fig. 7.4, the classification accuracy was set to 90% for both paradigms, while varying the number of repetitions within a trial. In Fig. 7.5, the number of repetitions was set to 5 and 6, while varying the classification accuracy (the average number of transitions in GIBS was defined as $N_{tr} = 0.60$). In Fig. 7.4, we see for example that it is possible to achieve a similar T_d by selecting 5 repetitions in RC, and 3 repetitions in GIBS (illustrated in Fig. 7.4 by the horizontal line segment). Therefore, to have the same T_d, the detection with GIBS would have to be achieved with two repetitions less than with the RC speller. Setting the number of repetitions and varying the accuracy, the curves in Fig. 7.5 show that, for example, for $N_{rep} = 5$, for an 80% accuracy in RC, a 90% accuracy is required in GIBS to achieve the same T_d, i.e., 10% more (illustrated in Fig. 7.5 by the horizontal line segment). In section 7.3, the two

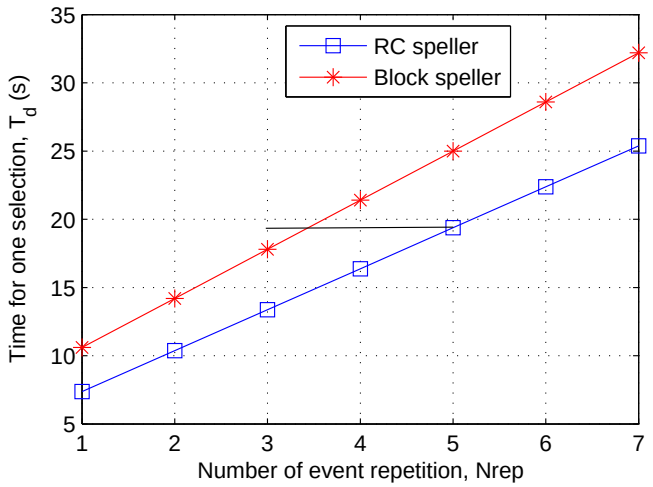


Figure 7.4: Time duration of a trial (T_d) for RC speller (7.3) and GIBS (7.4) varying the number of repetitions and setting the classification accuracy to 90%.

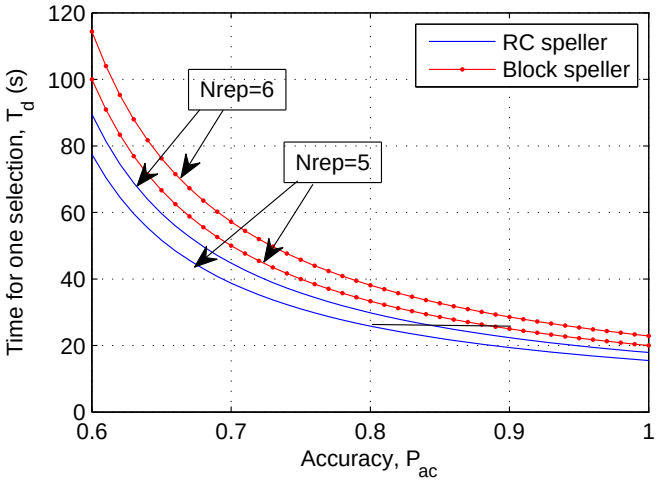


Figure 7.5: Time duration of a trial (T_d) for RC speller (7.3) and GIBS (7.4) varying the classification accuracy and setting the number of repetitions to 5 and 6.

spellers are compared online, inferring if GIBS can attain transfer rates similar to RC speller.

Table 7.1: *Case-study III*: summary of experimental protocol

Recording	
Channels:	12 Ag/Cl electrodes Fz, Cz, C3, C4, CPz, Pz, P3, P4, PO7, PO8, POz and Oz (10-20 extended positions) Reference: left or right ear lobe; Ground: AFz Vertical EOG and Horizontal EOG (eye monitoring)
Frequency sampling:	256 Hz
Preprocessing filters:	0.1-30 Hz bandpass filter and 50 Hz notch filter
Impedance:	$< 10K\Omega$
RC speller paradigm	
Paradigm:	6×6 matrix RC speller
RC parameters:	SOA=200 ms, ISI=100 ms, ITI=3.5 s
Screen:	15" at about 60-70 cm
—	
GIBS speller paradigm	
Paradigm:	GIBS speller
GIBS parameters:	SOA=150 ms, ISI=75 ms, ITI=3.5 s
Screen:	15" at about 60-70 cm
Participants	4 able-bodied

7.3 *Case-study III* - GIBS vs RCS

7.3.1 Experimental protocol

We assessed experimentally GIBS and compared its performance with that achieved with RC speller. The summary of experimental protocol is presented in Table 7.1. Two different experiments were performed. In one experiment, the participants were allowed to gaze the target symbols (overt attention), and in this second experiment they were not allowed to gaze the target symbols (covert attention). In the second experiment, vertical and horizontal electrooculographic signals (V-EOG and H-EOG) were recorded to monitor the movement of the eyes, ensuring that no ocular movements occurred during the covert attention task (in case of movement, the detection was discarded). Before the online session, several eye movement and blinks were recorded to serve as identification of these patterns. The online sessions were preceded by a calibration session of approximately 5 min. During the RC calibration phase, the user had to attend the letters of the word 'INTERFACE', and during GIBS calibration, the user had to attend 7 letters of the central group and 5 choices of small blocks. The labeled datasets obtained from calibration have 180 target epochs and 900 non-target epochs for the RC speller, and 96 target epochs and 1056 non-target epochs for GIBS. These datasets were used to obtain the models for online classification, through the C-FMS+qNB methodology presented in chapter 5.

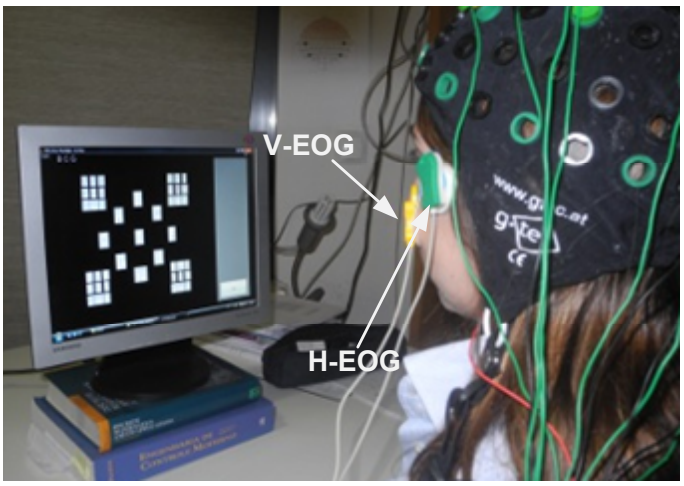


Figure 7.6: Subject controlling GIBS with covert attention. Experimental setup with H-EOG and V-EOG recording for eye movement monitoring.

7.3.2 Online results

Four participants tested both paradigms using overt attention. Two of them, S02 and S04, underwent the second experiment and tested both paradigms using covert attention. The participants had to spell the 19 characters sentence 'THE-QUICK-BROWN-FOX'. The results obtained with overt attention are in Table 7.2. The averaged results, using the number of repetitions that maximizes the PBR, show that the performance achieved with GIBS, regarding accuracy and number of repetitions, is substantially better than that obtained with RC. The results are 14.89 bpm, 85.5% accuracy and 4.75 repetitions for RC speller, and 16.67 bpm, 96.02 % accuracy and 3.5 repetitions for GIBS. The results show that the classification improvements with GIBS were enough to compensate the transitions between blocks, even reaching a PBR higher than that obtained with RC speller. Note however, that the performance of this group of participants was below the average results obtained with RC in *Case-study II*. As concerns the experiments based on covert attention, the online results are in Table 7.3. None of the participants was able to control the RC speller. Participants reported that they were unable to perceptually attend the targets placed 2 degrees outer the center (see Fig. 7.1d)). More than the decrease of the spatial acuity, the nearby distractors impeded target processing (crowding effect phenomenon) [Scolari 2007], a phenomenon that was augmented by the large size of flash events (entire row/column). On the other hand, both participants were able to covertly control GIBS with effective transfer rates. In GIBS paradigm, the crowding effect is reduced by the increased target-distractor distances, and by the small size of the flash

Table 7.2: Online results - **Overt** attention. Sentence spelled: 'THE-QUICK-BROWN-FOX'.

		RC speller		GIBS	
Subject	N_{rep}	P_{ac}(%)	PBR (bpm)	P_{ac} (%)	PBR (bpm)
S01	5	100	20.01	-	-
	3	68.4	10.68	100	19.64
	2	-	-	78.9	14.25
S02	5	78.9	11.58	100	13.98
	3	-	-	94.7	17.57
S03	5	78.9	11.58	100	13.98
S04	5	89.4	15.79	100	13.98
	4	84.2	16.20	-	-
	3	-	-	89.4	15.50
Average	4.75/3.5	85.5	14.89	96.02	16.67

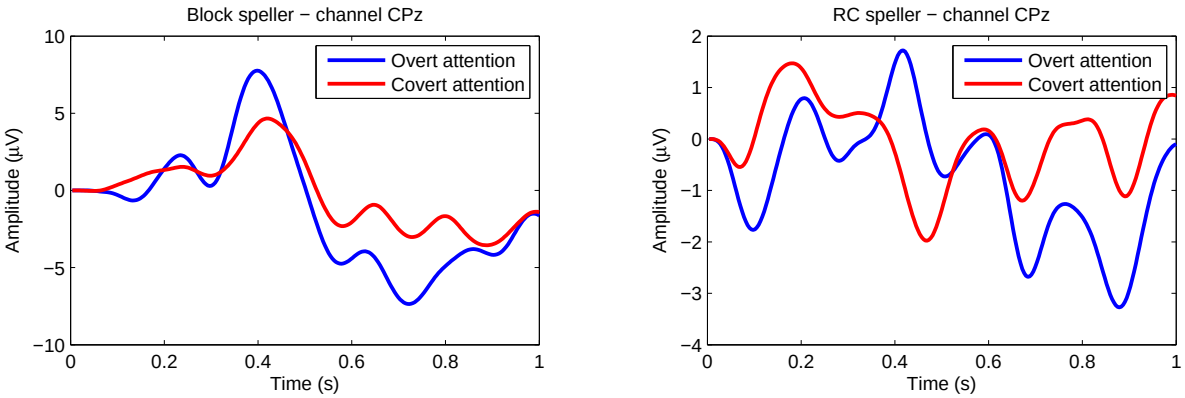
Table 7.3: Online results - **Covert** attention. Sentence spelled: 'THE-QUICK-BROWN-FOX'.

		RC speller		GIBS	
Subject	N_{rep}	P_{ac}(%)	PBR (bpm)	P_{ac} (%)	PBR (bpm)
S02	5	(a)	-	100	13.98
	3	-	-	84.2	13.43
S04	5	(a)	-	89.4	11.04

events. It is known that the crowding effect is eliminated entirely at a distance that is known as the critical spacing point [Scolari 2007]. In Tables 7.2 and 7.3, (a) means that participant was unable to perform the task, and (-) means that the experiment was not performed.

7.3.3 ERP waveforms

The average ERP waveforms elicited in overt and covert attention tasks were computed taking the data collected during the calibration phases. Fig. 7.7a) shows the waveforms of the ERP average for GIBS. Both overt and covert experiments evoked a P300, and as expected, the P300 with overt attention is higher. Fig. 7.7b) shows the waveforms of the ERP average for RC speller. The overt attention task elicits a P300, however there



a) b)

Figure 7.7: P300 average considering overt and covert task (datasets of participant S02). a) P300 waveforms for GIBS; b) P300 waveforms for RC speller.

is no traceable P300 in the covert attention task. The waveform shows only the strong oscillatory effect of SSVEP, typical of the RC speller (see chapter 5). These waveforms explain why participants were unable to control RC online.

7.4 Conclusion

Gaze-independent paradigms are the only choice for people with severe motor ocular restrictions. GIBS provided an efficient way to control P300-based BCIs without moving the eyes. While the validation of GIBS still requires further experiments with a larger number of subjects, it was clear that the RC speller does not allow covert control. The ineffectiveness of the RC paradigm has led many researchers, performing clinical experiments in patients with severe motor disability, to abandon this paradigm, and starting to investigate auditory paradigms.

Surprisingly, the transfer rate of GIBS outperformed the RC speller in overt control. These results support the importance of the paradigm design to increase BCI performance. While the control of GIBS is more complex, since it needs a two-level selection, its visual design, with fewer and larger symbols proved to be an important factor for BCI performance. Other approaches deserve to be investigated. For instance, the identification of symbols in the periphery can be improved by scaling magnification effects and increasing inter-symbol spacing. Mixed solutions might be worth investigating, i.e., a bit of covert attention but not full restriction of eye movements, adapted to each individual.

Chapter 8

Asynchronous Arrow-paradigm

Contents

"Sorry, I wasn't thinking."
A BCI participant, after spelling unintentionally a naughty word, when he was using a BCI without non-control state detection.

This chapter introduces a P300-BCI called Arrow-paradigm, designed specifically to control a robotic wheelchair. This BCI system was initially developed for synchronous control and then extended to asynchronous (self-paced) control. Two methods for asynchronous control were developed and analysed offline. One of them was implemented and assessed online.

8.1 Introduction

Most current BCI systems, including those previously presented in this thesis, work in synchronous mode, i.e., the system provides the temporal cues that define the time instants in which the user can input a command. In contrast, in asynchronous (self-paced) BCI, the user can decide when he/she wishes or not to control the interface, i.e., the user can switch between a 'control-state' and a 'non-control-state'. While this issue is important but not critical in P300 spellers, it is crucial when the BCI is intended to control a robotic device, such as a wheelchair. Without self-paced control, the BCI is always translating the EEG to commands. Thus, if the user does not want to provide any command, EEG patterns are translated to unintentional commands leading to false actions (false positives) with unpredictable consequences. Therefore, asynchronous control is highly desirable for effective BCI in certain real-world scenarios. This issue has been addressed in [Zhang 2008] for P300-based BCI, and in [Mason 2000, Townsend 2004, Scherer 2004] for SMR-based BCI.

The implementation of asynchronous control in P300 BCIs requires the detection of a non-control state, for which no commands are sent by the interface. Two approaches will be investigated and assessed in the context of a paradigm designed to send commands to a robotic wheelchair, henceforth designated by Arrow-paradigm.

8.2 *Case-study IV* - Asynchronous and synchronous control of Arrow-paradigm

8.2.1 Arrow-paradigm design

The Arrow-paradigm was initially designed to be operated in synchronous mode [Pires 2008b]. A screenshot is shown in Fig. 8.1. The paradigm provides a small lexicon which includes assistance words and information to steer a mobile robot (e.g., a wheelchair), namely, 'GO', 'LEFT', 'RIGHT', 'STOP', 'YES', 'NO', 'HELP', 'PHONE', 'FOOD', 'WATER' and 'ROOM' (or the equivalent words in portuguese). Each symbol/word is randomly intensified with 175 ms SOA and 75 ms ISI. When the symbol/word is intensified , the color is changed and the size is slightly increased. The number of symbols is 11, leading to target probability of 1:11. The inter-trial interval was settled to 2.5 seconds. This synchronous paradigm was successfully tested online by several able-bodied and disable participants. Given its simplicity, presenting larger and fewer symbols

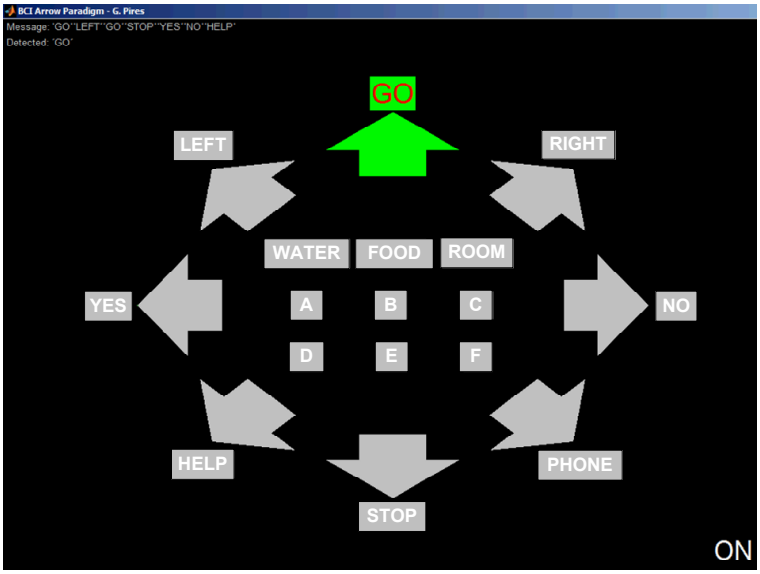


Figure 8.1: Screenshot of the Arrow-paradigm. It provides a set of direction commands and a small lexicon composed by assistance words. Symbols $\{A \cdots F\}$ are not currently used, simply indicating possible future commands.

than RC and LSC spellers, most of the participants achieve higher performance than that achieved with those spellers. Some participants are actually able to control this paradigm with single trial (one repetition). However, this paradigm is of little use, if the user is not able to decide when he/she wants to send a command. Initially, we successfully tested some simple approaches to switch ON/OFF the system, by, for example, closing the eyes, which increases the amplitude of the α rhythm, or by gazing a flash at a given frequency, which induces a SSVEP with the same flashing frequency. The switch could be activated by the detection of these signals. However, these approaches require strong ocular movements and therefore they have no relevance if the system is to be used by individuals with severe motor disabilities. The alternative is to detect a non-control state relying only on EEG without resorting to any type of muscular activity.

8.2.2 Experimental setup and calibration protocol for asynchronous control

A set of online experiments were performed using the Arrow-paradigm in synchronous and asynchronous mode. The summary of the experimental protocol is presented in Table 8.1. Six able-bodied (S01-S06) and four disabled participants controlled the Arrow-paradigm in synchronous mode. The motor disabled participants are the subjects identified as S19,

Table 8.1: *Case-study IV*: summary of experimental protocol

Recording	
Channels:	12 Ag/Cl electrodes Fz, Cz, C3, C4, CPz, Pz, P3, P4, PO7, PO8, POz and Oz (10-20 extended positions) Reference: left or right ear lobe; Ground: AFz
Frequency sampling:	256 Hz
Preprocessing filters:	0.1-60 Hz bandpass filter and 50 Hz notch filter
Impedance:	$< 10K\Omega$
Synchronous Arrow-paradigm	
Parameters:	SOA=175 ms, ISI=75 ms, ITI=2.5 s
Screen:	15" at about 60-70 cm
Participants	6 able-bodied and 4 motor disabled
—	
Asynchronous Arrow-paradigm	
Parameters:	SOA=175 ms, ISI=75 ms, ITI=2.5 s
Screen:	15" at about 60-70 cm
Participants	2 able-bodied

S20, S23 and S24 in Table 6.2, who tested previously RC and LSC paradigms in *Case-study II*. They are all confined to a wheelchair, and thus they could all benefit from this interface.

The asynchronous mode was tested only by able-bodied subjects S01 and S02. To investigate the non-control state, the standard calibration protocol procedure introduced in previous chapters was changed in order to accommodate the additional non-control (NC) state in between control (C) states. In the control state, participants were asked to attend the target events indicated at the top of the screen, exactly as it was done in previous synchronous paradigms. In the non-control state, participants were asked to relax, not paying attention to any of the events (idle state). At the top of the screen, a message 'non-control - RELAX' is provided to the user. During this period, the events continue flashing, and the user keeps looking[1] to the screen, simulating a real online scenario. The calibration session is divided in the succession of states: C + NC + C + NC + C. During the control state, user attends successively three target symbols, each one corresponding to a round of 10 repetitions of the events. The non-control state corresponds to the time equivalent to one symbol round. At the end of the calibration session, a dataset contains 90 target epochs, 900 non-target epochs and 220 non-control epochs, and so we have now three classes instead of two.

[1]This is the worst case scenario, considering that the user can not move the head or the eyes to look away.

8.2.3 Three-condition data analysis

Using the datasets recorded in the asynchronous calibration, we searched for discriminative EEG features between the 3-classes: target, non-target and non-control. We started by analysing the three conditions in the time domain, computing the average of the epochs for each class, as shown in Fig. 8.2a) for a given dataset of subject S01. Both non-target and non-control epochs evidence the SSVEP effect of the flashes, at a frequency $f = 5.7$ Hz ($5.7 = 1/0.175$, SOA=175 ms). Non-control epochs show additional oscillatory components identified in the frequency domain. Fig. 8.2b) shows the frequency spectrum of the three classes, obtained by computing the FFT for each epoch and then averaging the FFTs. Dividing the epochs in control state {target + non-target} and {non-control} state, we obtained the plots in Fig. 8.2c)d). Some discriminative frequencies between control and non-control state seem to occur at $\approx$ 6, 17 and 34 Hz, however they varied across sessions and across subjects.

Based on the above analysis, two approaches were investigated. The first one to detect {target + non-target} vs. {non-control} and the second one to detect {target} vs. {non-target + non-control}. In the first approach, we investigated if it was possible to detect discriminative features related to stimuli frequencies, or frequencies related to the relaxing condition of the non-control state. For instance, α rhythms are associated with idle and meditation states while β rhythms are associated to sustained attention [Monastra 2005]. To implement the {control-state} vs. {non-control state}, discriminative features need to be selected, automatically adjusting to each individual. Therefore, we extended the spatio-spectral filter SSFCB (5.27) described in section 5.3.4 to automatically tune the frequency shaping filter. This method, henceforth called ASSFCB (automatic SSFCB) is described in section 8.2.4. The second approach implements a detector between {target} vs. {non-target + non-control}. Non-control epochs share many similarities with non-target epochs, thus it is expected to obtain classification rates that approach those obtained between target vs. non-target. The non-control detection could be attained by imposing a threshold limit in the posterior probabilities [Zhang 2008], controlling thereby the rate of false positives (FPR). This approach is described in section 8.2.5.

8.2.4 ASSFCB spatial filter - detection of non-control state

In section 5.3.4, we presented SSFCB, a spatio-spectral filter. The weights of the spectral filter $\tilde{\mathbf{h}} = [\beta(1) \quad \beta(2) \quad \cdots \quad \beta(T')]$ were tuned manually. We propose here ASSFCB, an extension of SSFCB that allows the automatic tuning of the spectral weights.

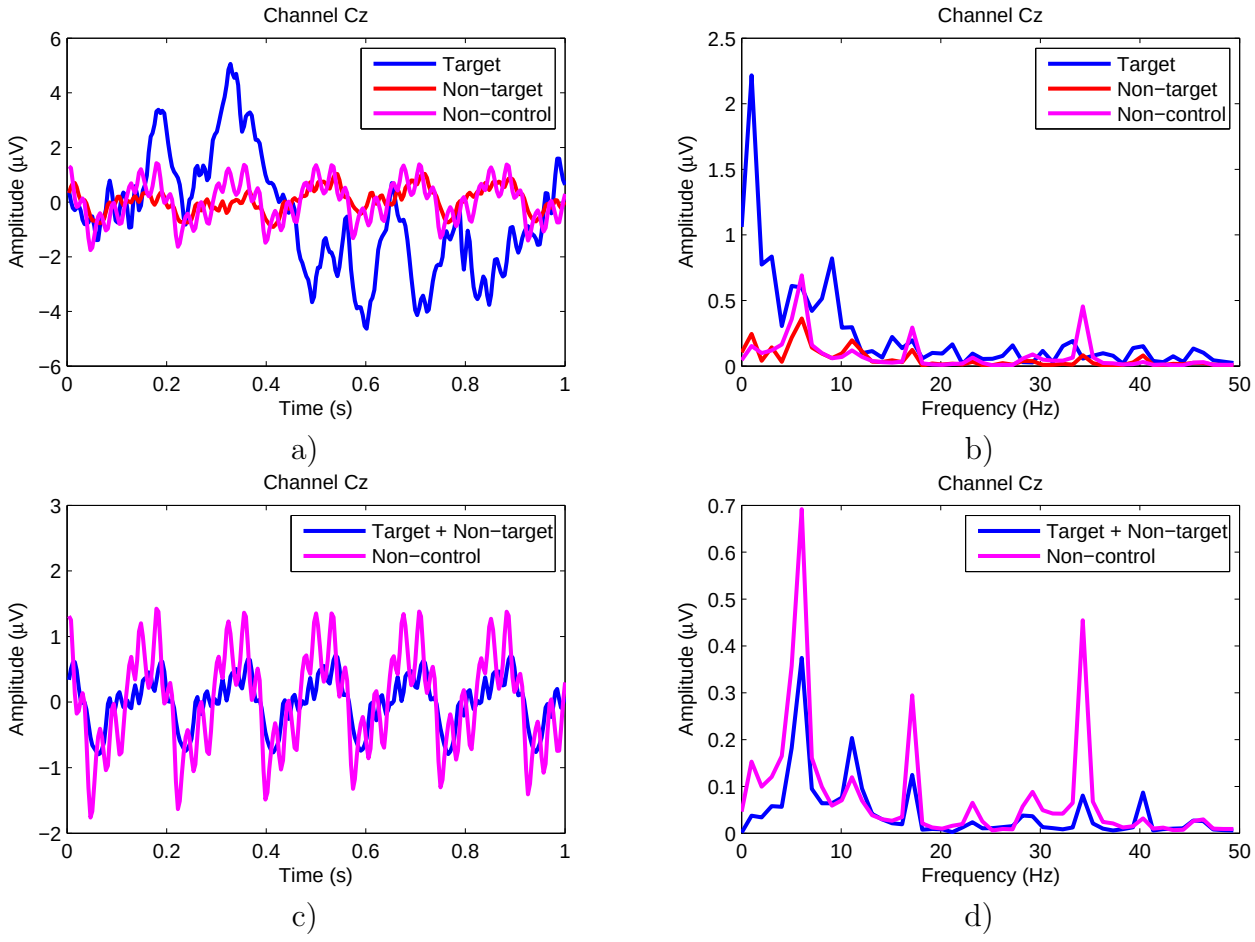


Figure 8.2: Dataset from a representative subject (90 target epochs, 900 non-target epochs and 220 non-control epochs) a) Average in the time domain of target, non-target and non-control epochs; b) Spectrum corresponding to the average of the FFTs of the epochs taken for (a); c) Average in the time domain of {target + non-target} and non-control epochs; d) Spectrum corresponding to the average of the FFTs of the epochs taken for (c).

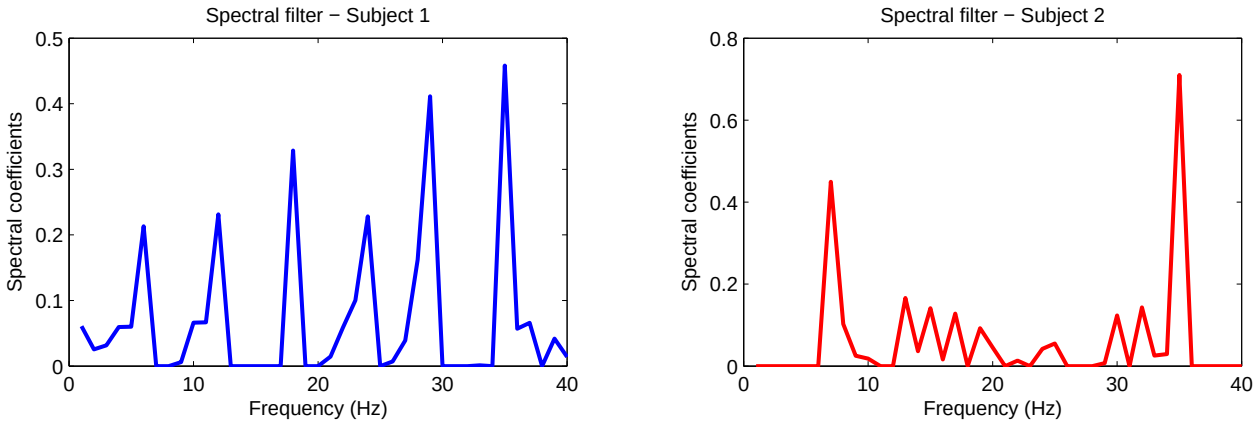
Some automatic approaches combining CSP spatial filter and temporal/spectral filters were already proposed for motor-imagery-based BCIs [Lemm 2005, Dornhege 2006, Tomioka 2006, Wu 2008]. Our approach follows the same underlying concepts used in [Tomioka 2006, Wu 2008]. ASSFCB is detailed in Algorithm 1. The method is fully implemented in the frequency domain following the mathematical formalism introduced in section 5.3.4. The algorithm implements an iterative procedure that finds the spatio-spectral filters in two separate steps, spatial filter and spectral filter, following a sub-optimum approach. The spectral filter $\tilde{\mathbf{h}}$ is initialized with all coefficients set to 1, corresponding to an all-pass filter. In the first iteration, the spatial filter $\tilde{W}$ is estimated taking $\tilde{\mathbf{h}} = [1 \quad 1 \cdots 1]$ and thus it coincides with the original FC beamformer (FCB) filter. In

the subsequent step, the spatio-spectral projections $\tilde{\mathbf{y}}$ are used to estimate the optimal spectral filter $\tilde{\mathbf{h}}$, through a frequency domain eigenfilter [Haykin 1996]. Combining the two filters, spatio-spectral projections are computed and transformed to the temporal domain, and then qNB classifier infers the performance of spatio-spectral filters, measuring the classification error rate. The iterative procedure is repeated until a maximum number of iterations is reached or until the classification error rate is below a predefined threshold.

ASSFCB was used to discriminate between {control-state} vs. {non-control-state}. We built the control-state joining target epochs and non-target epochs $\{C = T+NT\}$, and the non-control state was built from relaxing/idle state epochs {NC}. As we saw before, non-target and non-control state epochs are very similar, which makes this detection a challenging problem. The optimum spectral filters learned by ASSFCB are shown in Fig. 8.3 for subjects S01-S02. The spectral coefficients are zero for low frequencies, which indicates no discrimination for frequencies in the range of target ERPs. Some high

Algorithm 1 ASSFCB

1: Spectral filter initialization:
$\tilde{\mathbf{h}} = [\beta(1) \quad \beta(2) \quad \cdots \quad \beta(T')]$ with $\beta(j) = 1, \quad j = 1 \cdots T'$
2: Obtain Discrete Fourier Transform for each epoch k in dataset
3: **for** $k = 1 : K$ **do**
4: $\tilde{\mathbf{X}}^{(k)} = \mathrm{DFT}(\mathbf{X}^{(k)}), \quad \tilde{\mathbf{X}} \in \mathbb{R}^{N \times T'}$
5: **end for**
6: **repeat**
7: Obtain spectral filter projections according to (5.26):
$\tilde{\mathbf{Z}}^{(c)} = \tilde{\mathbf{X}}^{(c)}.\tilde{\mathbf{h}} \quad (c \in \{+, -\})$
8: Obtain spatial covariances $\tilde{\mathbf{S}}_b$ and $\tilde{\mathbf{S}}_w$ and then spatial filters according to maximization of (5.27):
$\max \frac{\tilde{W}^H \tilde{\mathbf{S}}_b \tilde{W}}{\tilde{W}^H \tilde{\mathbf{S}}_w \tilde{W}}$
Select spatial filter $\tilde{W}^{(1)}$ associated to top eigenvalue
9: Compute spatio-spectral filter projections according to (5.28):
$\tilde{\mathbf{y}}^{(c)} = \tilde{W}^{(1)\prime} \tilde{\mathbf{Z}}^{(c)}$
10: Compute classification error rate applying classifier qNB to spatio-spectral projections in the time domain
$\mathbf{y}^{(c)} = \mathrm{IDFT}(\tilde{\mathbf{y}}^{(c)})$
11: Compute the spectral filter $\tilde{\mathbf{h}}$ in the frequency domain (eigenfilter) by maximizing:
$\max \frac{\tilde{h}^H \tilde{\mathbf{R}}_b \tilde{h}}{\tilde{h}^H \tilde{\mathbf{R}}_w \tilde{h}}$, s.t. $\beta(j) \geq 0$, where $\tilde{\mathbf{R}}_b$ and $\tilde{\mathbf{R}}_w$ are the estimated spectral between-class and within-class covariances
12: Normalize the spectral filter $\tilde{\mathbf{h}}$ according to:
$\tilde{\mathbf{h}} = \frac{[\beta(1) \quad \beta(2) \quad \cdots \quad \beta(T')]}{\|[\beta(1) \quad \beta(2) \quad \cdots \quad \beta(T')]\|_2}$
13: **until** Max number of repetitions reached or error rate $<$ threshold
14: Select spatio-spectral filter that minimizes the classification error rate



a) b)

Figure 8.3: Optimal spectral filters $\tilde{\mathbf{h}}$ obtained for subject 1 (a) and subject 2 (b) learned by ASSFCB.

coefficients (spectral peaks) seem related to SSVEP evoked by stimuli frequencies. These frequencies overlap rhythms in the α band, which makes difficult to infer whether spectral coefficients are related to attention or relaxation states. Curiously, a high discriminative frequency around 35 Hz is detected by ASSFCB in both subjects. This deserves a future neurophysiological analysis. To assess the performance of ASSFCB across iterations, we computed the classification error rate for each iteration as shown in Fig. 8.4. Note that the first iteration coincides with the original spatial filter. The subsequent iterations show a significative improvement of the classification, showing the effectiveness of ASSFCB. This improvement is reached just after two or three iterations, and subsequent iterations may reduce the performance. Despite this significative improvement, the classification results were not sufficiently robust for online application.

8.2.5 Posterior probabilities - threshold adjustment

This approach takes the values returned from the posterior probabilities $P(C_i|\mathbf{y})$ computed from (5.30) using the Bayes theorem. Let us consider the three classes C_i with $i \in \{T, NT, NC\}$, then we have $P(C_T|\mathbf{y})$, $P(C_{NT}|\mathbf{y})$ and $P(C_{NC}|\mathbf{y})$ respectively for target, non-target and non-control. These probabilities were computed applying the qNB classifier to spatial FCB projections. The log-transformation was applied to (5.29), obtaining

$$P(C_i|\mathbf{y}) = \log(\pi_i) - \sum_{j}^{N_f} \log(\sigma_i(j)) - \frac{1}{2}\sum_{j}^{N_f} \left((\mathbf{y}(j) - \mu_i(j))\sigma_i^{-2}(j)(\mathbf{y}(j) - \mu_i(j))\right). \quad (8.1)$$

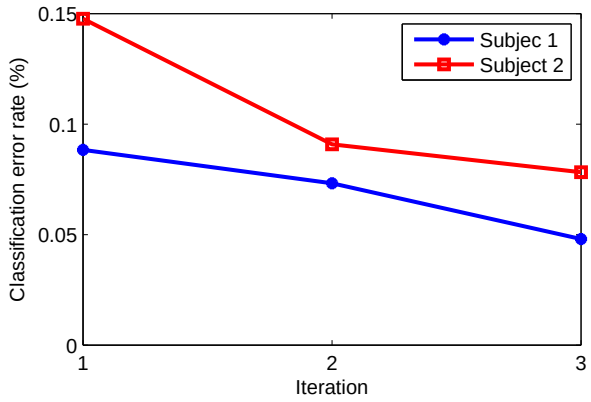


Figure 8.4: Classification performance (error rate) of ASSFCB across three iterations. The first iteration coincides with the original FCB spatial filter.

For each binary classifier presented in Fig. 8.5, the posterior probabilities were normalized according to

$$\bar{P}(C_i|\mathbf{y}) = \frac{P(C_i|\mathbf{y})}{\sum_i P(C_i|\mathbf{y})}. \tag{8.2}$$

The right and left distributions in Fig. 8.5a) correspond respectively to

$$\bar{P}(C_T|\mathbf{y}_T) \wedge \bar{P}(C_{NT}|\mathbf{y}_{NT}) \tag{8.3}$$

and

$$\bar{P}(C_{NT}|\mathbf{y}_T) \wedge \bar{P}(C_T|\mathbf{y}_{NT}) \tag{8.4}$$

i.e., (8.3) and (8.4) give respectively the distributions of posterior probabilities for correct and false detections. The same applies for each of the others binary classifications in Fig. 8.5 b)c)d). Simulating the real scenario, we tested a binary classifier {target} vs. {non-target + non-control} with two different datasets, in Fig. 8.5c) and 8.5d). The posterior probabilities show that the classifier performs only slightly worse than {target} vs. {non-target} binary classifier. Adjusting a threshold Δ, we can detect the control state by eliminating the FPR. The online implementation of asynchronous control requires a simple modification of (5.36) in order to include this threshold. Rule (5.36) assumes that a P300 ERP is always present, while in asynchronous mode, P300 may not exist. Therefore, taking the events $j \in \{1, \cdots, 11\}$, the detection of a target symbol is obtained

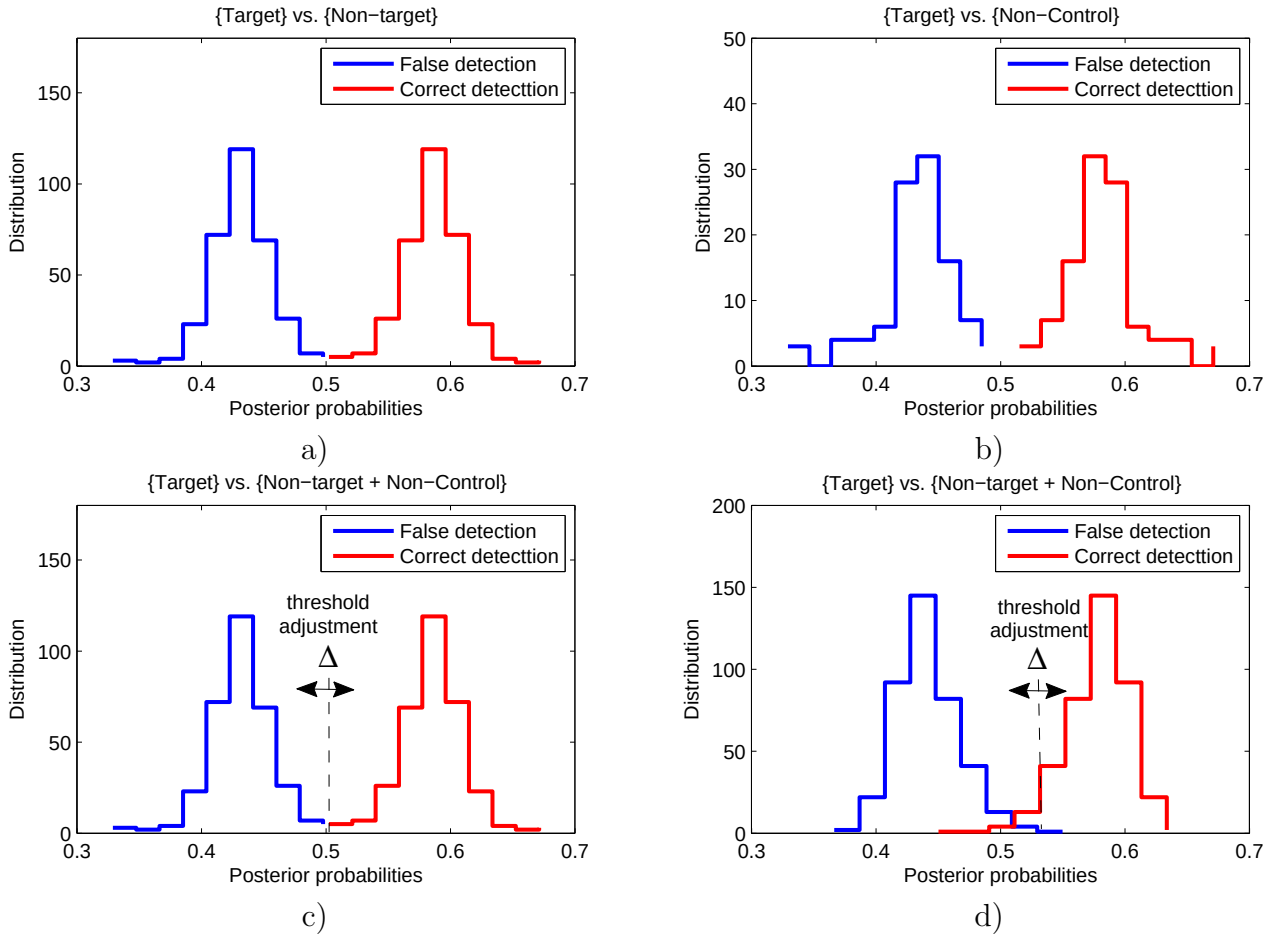


Figure 8.5: Distribution of posterior probabilities computed from (8.2) using (FCB + qNB). Three binary classifiers were tested. a) {target} vs. {non-target}; b) {target vs. non-control}; c) {target} vs. {non-target + non-control}; d) Same as c) but with a different dataset. All graphs were obtained using a K-epoch average, $K = 3$.

according to

$$\begin{cases} \text{if } \exists \bar{P}^j(C_T|\mathbf{y}) > \Delta & \text{then Target} = \underset{j\in\{1,\cdots,11\}}{\arg\max}\, \bar{P}^j(C_T|\mathbf{y}) \\ \text{else} & \text{Non-control state} \end{cases} \tag{8.5}$$

8.2.6 Online results

A set of online experiments were performed using the Arrow-paradigm in synchronous and asynchronous mode according to protocol defined in Table 8.1.

8.2.6.1 Synchronous control

For synchronous control, participants performed a standard calibration session. Then, simulating the online steering of a wheelchair, they were asked to send the following 15 symbol/command sequence: 'GO', 'LEFT', 'GO', 'STOP', 'YES', 'NO', 'HELP', 'FOOD', 'WATER', 'PHONE', 'GO', 'RIGHT', 'STOP', 'ROOM' and 'STOP' (or equivalent words in portuguese). The users were seated at a standard chair or on their own wheelchairs. Table 8.2 summarizes the results, showing for each participant, the number of repetitions N_{rep} for which the minimum error rate was achieved. Since we want to minimize the number of false steering commands, the error rate metric plays an increased relevance to evaluate the BCI performance. The number of symbols per minute was computed taking the ITI time, simulating a realistic scenario. Several able-bodied participants achieved a 7.25 SPM with 100% accuracy. Motor disabled participants achieved 4 to 5 SPM, but with some errors. Nevertheless, these subjects performed better with this paradigm than with RC and LSC paradigms (compare with Table 6.6), which is mainly explained by the higher simplicity of Arrow-paradigm.

8.2.6.2 Asynchronous control

After a calibration session, following the protocol described in section 8.2.2, subjects S01-S02 controlled asynchronously the Arrow-paradigm. The classification methodology was the one described in section 8.2.5. The threshold Δ was adjusted offline from calibration data to minimize the FPR (non-control states detected as target symbols) while trying not increasing to much the rate of false negatives, FNR (target symbols detected as non-control states). The first experiment consisted of a sequence of intermittent symbol detection and non-control state as shown in Fig. 8.6a), comprising a total of 36 symbol detections plus 35 non-control state detections (sequence type I). A FPR of 11.4% and 5.5% was achieved respectively by subjects S01 and S02, while keeping a low FNR. A second type of sequence was tested, consisting of 34 non-control states and 12 symbols detection (sequence type II in Fig. 8.6b)). The sequence intended to test a scenario in which the user is mainly in non-control state and sends a command occasionally. An FPR ranging from 3-6% was attained also with a low FNR. These quite good results were achieved for a relatively high SPM (7.25 SPM, $N_{rep} = 3$). To improve the reliability of the BCI system, we could further reduce the FPR. This can be reached by increasing N_{rep}, which decreases the SPM, or by increasing the threshold level Δ, which increases the FNR. Naturally, reliability and SPM are inversely proportional.

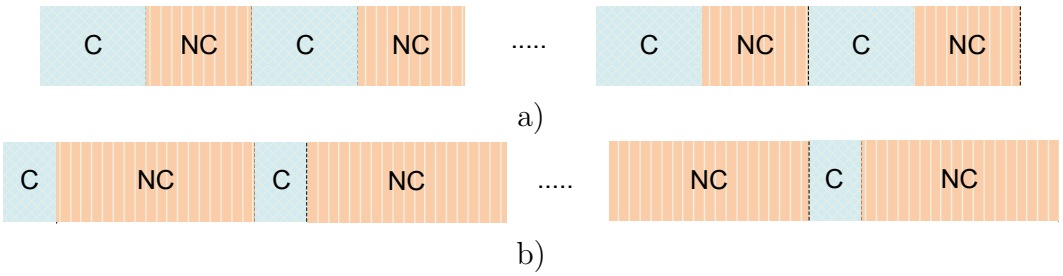


Figure 8.6: Sequences of control an non-control states: a) sequence type I - intermittent control (C) and non-control (NC) sates; b) sequence type II - non-control (NC) states with occasional control (C) states.

Table 8.2: Online results for Synchronous control of Arrow-paradigm. Detection of a 15 symbol/command sequence. Clinical data of subjects S19, S20, S23, S24 are described in Table 6.2

Synchronous Arrow-paradigm			
Subject	N_{rep}	Nr. errors (%)	SPM with ITI
S01	3	0 (0%)	7.25
S02	3	0 (0%)	7.25
S03	5	0 (0%)	4.95
S04	5	0 (0%)	4.95
S05	3	0 (0%)	7.25
S06	5	4 (27%)	4.95
S19	(a)	-	-
S20	5	3 (20%)	4.95
S23	5	2 (13%)	4.95
S24	6	2 (13%)	4.27

Table 8.3: Online results for Asynchronous control of Arrow-paradigm: 36 symbol detection with 35 non-control states in between (sequence type I).

Asynchronous Arrow-paradigm					
Subject	N_{rep}	Errors (%)	False Positives (%)	False Negatives (%)	SPM with ITI
S01	3	1 (2.7%)	4 (11.4%)	0 (0%)	7.25
S02	4	1 (2.7%)	1 (2.8%)	1 (2.7%)	5.88
	3	0 (0%)	2 (5.5%)	2 (5.7%)	7.25

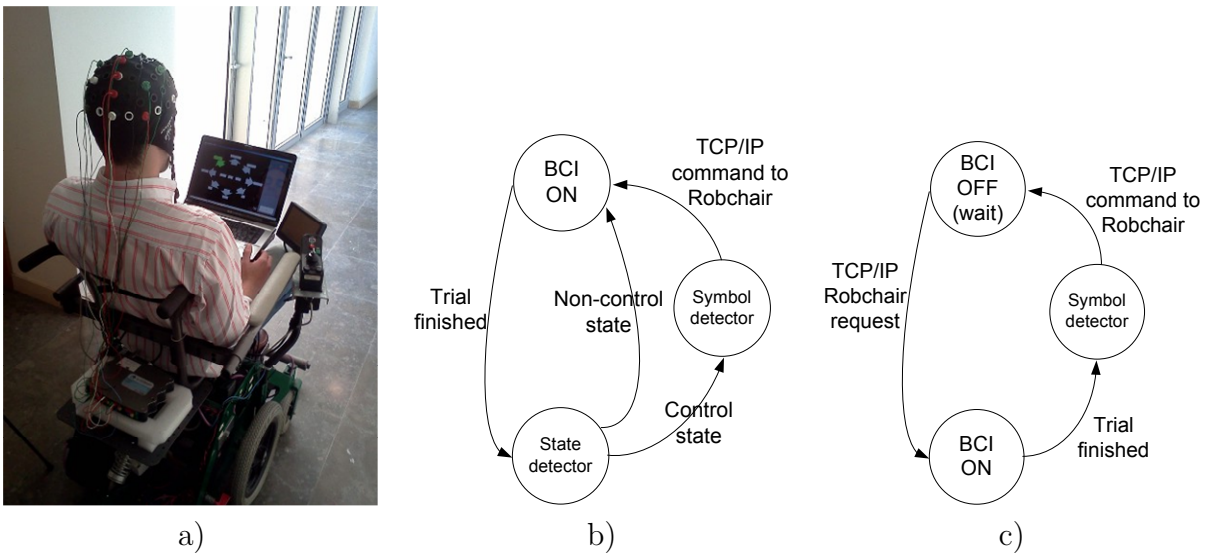
8.2.7 Wheelchair steering scenario

The number of decoded symbols per minute (SPM) in a BCI is very low, which means that users can only provide a few discrete commands per minute. Thus, the control of

Table 8.4: Online results for Asynchronous control of Arrow-paradigm: 34 non-control states with 12 symbols detection in between (sequence type II).

Asynchronous Arrow-paradigm					
Subject	N_{rep}	Errors (%)	False Positives (%)	False Negatives (%)	SPM with ITI
S02	3	0 (0%)	1 (2.9%)	1 (8.3%)	7.25
S02	3	0 (0%)	2 (5.8%)	0 (0%)	7.25

a wheelchair must rely on a navigation system that receives sparse commands from the user and that performs safe and smooth manoeuvres according to steering information [Pires 2002]. The navigation and human-machine shared control is currently being investigated in our Lab within our research group [Lopes 2011]. Several brain-actuated wheelchairs have already been proposed in other Labs. Motor imagination was investigated in [Galán 2008] to detect 3 mental steering commands (forward, left and right) and in [Tanaka 2005] to detect left and right mental states. The P300 ERP control has been investigated in [Iturrate 2009] and [Rebsamen 2010], the last one also addressing non-control state. In [Rebsamen 2010] the BCI allows the selection of high-level pre-defined locations (e.g. office) while in [Iturrate 2009] the BCI selects low-level locations in a 3D map of the surrounding environment. These systems are currently being developed as a proof-of-concept. As the number of decoded symbols is still very limited, the main



a) b) c)

Figure 8.7: a) Robchair prototype during an experiment at ISR; b) and c): state diagrams implemented in BCI to manage TCP/IP requests and commands between BCI and RobChair in operation mode 1 (b) and operation mode 2 (c).

challenge of a brain-controlled wheelchair is from the navigation side.

Some preliminary and pilot experiments were performed with the Robchair prototype (Fig. 8.7a)). A TCP/IP communication module was incorporated in the BCI system to send commands and receive requests from the wheelchair. At the present time, Robchair has two operation modes. One in which the user sends commands whenever he/she wants, and the other in which the user responds only to requests of Robchair in ambiguous scenarios (e.g., decide between left or right in bifurcations). These two operation modes are implemented in the BCI system based on state diagrams, as shown in Fig. 8.7b) and Fig. 8.7c) respectively. The description of the navigation experiments are beyond the scope of this thesis.

8.3 Conclusion

Asynchronous control is an important issue for effective use of BCI in real-world scenarios. Comparing the two approaches proposed for asynchronous control, the detection between {Target} vs. {Non-target + Non-control}, using the information from posterior probabilities, was more efficient than the detection between {Target + Non-target} vs. {Non-control}, using ASSFCB. This is explained by the high similarities between Non-target and Non-control epochs, both mainly reflecting the stimuli frequencies. Thus, the ASSFCB was unable to identify the idle/relax state during non-control, which may suggest that users may have to train this state condition. This approach requires further analysis. On the other hand, the approach detecting {Target} vs. {Non-target + Non-control} was quite efficient. Although the online results have been obtained only for two individuals, the number of experiments was extensive, validating the approach.

Chapter 9

Final remarks and conclusion

Contents

"One should not pursue goals that are easily achieved. One must develop an instinct for what one can just barely achieve through one's greatest efforts."
Albert Einstein (Nobel prize in Physics in 1921)

Our proposed work represents an effort to go one step forward to the application of P300-based BCIs in real-world scenarios. Nonetheless, many of the BCI issues that were investigated can be further improved for a more effective clinical and non-clinical application. This chapter reminds some of the main contributions and draws some of our lines of future research. Two satellite applications, developed in parallel with the main research work are also succinctly described in appendices C and D, namely a BCI-Tetris [Pires 2011a], and an automatic detection of P300 modulated by saliency of visual target [Teixeira 2011].

9.1 Increasing transfer rates of P300-based BCIs

Increasing the communication rate of BCI is one the most important issues for the effective application of BCI. The performance results achieved in chapters 5 and 6 were above the state-of-the-art, and therefore extremely motivating. Nevertheless, several of the proposed methodologies can be further improved to increase the transfer rates of P300-based BCIs, some of them guiding our future research, namely:

- Results obtained in this thesis confirmed that the design of the visual paradigm has a strong direct effect on user's performance. Paradigms using fewer symbols, such as GIBS and Arrow-paradigm, evoke enhanced P300 components, leading to higher classification accuracies, than those obtained with paradigms with a greater number of symbols such as RC and LSC spellers. As discussed in chapters 6 and 7, the P300 evoked by the standard RC paradigm is adversely affected by many undesirable factors. Although the decrease in the number of symbols of a paradigm enhances the P300 ERP, it happens that reduces the ITR. Therefore, more research should be carried out, trying to develop paradigms that avoid undesirable effects on P300, while keeping an effective number of symbols. This can be achieved by exploring new layouts and event strategies, as it was done in LSC speller. Effective event strategies can as well reduce the overall time of a trial. LSC can potentially further decrease the trial time, by extending the left/right event strategy to the four quadrants of the screen;

- Visual paradigms should be designed to evoke new discriminative neurophysiologic features, thereby introducing new discriminative dimensions. In this regard, the proposed LSC paradigm has several aspects that can be explored in the future. For instance, the left and right asymmetries of P300 targets related to visuo-spatial attention, and the phase shift evoked by left and right event strategies, can provide new left/right discriminative features, potentially increasing classification accuracy;

- Signal processing (feature extraction) is one of the most important components of a BCI. Most of current feature extractors applied in P300-based BCIs rely on amplitudes collected over time. Our analysis showed that frequency and spatial information provide important features for P300 detection. New methods should be investigated, combining temporal, frequency, phase and spatial neurophysiologic features.

9.2 Key issues for effective application of P300-based BCIs

In addition to communication rates, two other key issues, very important to the effective deployment of P300-based BCI in real-world scenarios, were addressed in this thesis. Chapter 7 introduced GIBS, an efficient gaze-independent speller, that ensures an effective BCI control by people with severe motor ocular limitations. The second key issue,

the asynchronous control (self-paced) described in chapter 8, is of crucial importance to control physical devices. One of the proposed approaches has proved extremely effective, providing an efficient mechanism to control the rate of false positive commands.

Unsupervised adaptation was not addressed in this thesis, but still remains a key challenge for BCI [Krusienski 2011]. Due to psycho-neurophysiological factors, neuroplasticity, and electrode setup, subjects exhibit an EEG variability from session to session or even within the same session. Most of P300-based systems perform a calibration session before each online operation to fit the classification model to the most recent condition of the user. Even minimizing the calibration time as it was done in our systems, this calibration is always a constraint to BCI use. After a single initial calibration, an efficient unsupervised adaptation could take the initial models adjusted to the user, and adapt them according to the current condition of the user. Adaptation can be implemented in different domains such as feature extraction, models, classification or post-processing [Krusienski 2011].

9.3 Conclusion

In this thesis, we described several P300-based systems with important contributions to signal processing and visual paradigm design. In the signal processing field, we proposed several novel statistical spatial filters in the time domain, which were then extended to the frequency domain, leading to new statistical spatio-spectral filters. These feature extractors provided significative classification improvements independently of the classifier type. Moreover, the feature extraction and classification methodology is computationally very fast and needs only a small amount of training data, thereby contributing to a small calibration time. In regard to paradigm design, LSC, a lateral single character speller, was designed to overcome several limitations of the standard RC speller, and to explore layout, event strategy, and hemispheric asymmetries in visual perception to improve the performance of BCI. Performance results obtained with LSC were above those obtained with RC, thus validating the design of the paradigm. LSC still has a margin for potential improvement in bit rate and accuracy. Spatial filters and LSC were thoroughly validated online and offline by extensive sets of experiments performed by able-bodied and individuals with motor disabilities. Performance results obtained by able-bodied were above state-of-the-art. ALS participants, in early and intermediate stages of the disease, performed only slightly below able-bodied participants. The performance of individuals with CP was lower, but two of them controlled the BCI effectively. Two

additional paradigms have been implemented with particular relevance for deployment of BCI in real-world scenarios, namely, GIBS, a gaze independent speller, and Arrow-paradigm, an asynchronous BCI, specially designed to steer a robotic wheelchair. These two paradigms were validated online by able-bodied subjects.

We have worked toward the effective application of BCI, but much can still be done in many areas to improve its application. The use of BCI goes far beyond clinical application and will be part, certainly, of our daily life in the near future. Neuroplasticity will help us to adapt ourselves to better control BCI, and BCI will induce new processes of neuroplasticity, a dichotomic relation that may contribute to an optimized use of all the brain potential abilities.

Appendix A

Bit rate metrics

A.1 Derivation of the channel capacity of a BCI

The channel capacity is obtained from the mutual information of the channel [Shannon 1948]

$$I(X;Y) = H(X) - H(X|Y) \tag{A.1}$$

where $H(X)$ is the source entropy and $H(X|Y)$ is the information lost in the noisy channel. Assuming N_s possible equiprobable choices (probability $P(x_i) = 1/N_s$) of the BCI paradigm, and the conditional probability $P(x_i|y_i)$ given by the BCI accuracy, p_a, then $I(X;Y)$ is given by

$$\begin{aligned} I(X;Y) &= \sum_x P(x_i)\log\frac{1}{P(x_i)} - \sum_{x,y} P(x_i y_j)\log\frac{1}{P(x_i|y_j)} \\ &= \sum_x P(x_i)\log\frac{1}{P(x_i)} - \sum_{x,y} P(x_i|y_j)P(y_j)\log\frac{1}{P(x_i|y_j)} \\ &= \sum_x P(x_i)\log\frac{1}{P(x_i)} - [P(x_1|y_1)P(y_1)\log\frac{1}{P(x_1|y_1)} + \cdots + \\ &\qquad + \cdots + P(x_1|y_{N_s})P(y_{N_s})\log\frac{1}{P(x_1|y_{N_s})}]N_s \\ &= \log(N_s) - [p_a\log\frac{1}{p_a} + (1-p_a)\log\frac{1}{\frac{1-p_a}{N_s-1}}] \\ &= \log(N_s) + p_a\log(p_a) + (1-p_a)\log\frac{1-p_a}{N_s-1}. \end{aligned} \tag{A.2}$$

The channel capacity $I(X;Y) \equiv B$, represents the amount of information transferred per symbol, measured in bits/symbol.

Appendix B

CSP and generalized eigenvalue spatial filter

B.1 Common spatial patterns

Consider the $N \times N$ EEG spatial covariance normalized as $\mathbf{R} = \mathbf{X}\mathbf{X}'/tr(\mathbf{X}\mathbf{X}')$ where $'$ represents the transpose operator and $tr(X)$ represents the trace of X. The normalized spatial covariance is computed separately for class-1 and class-2, respectively $\mathbf{R_1}$ and $\mathbf{R_2}$. The CSP filter as introduced in [Soong 1995] consists in finding a matrix W and a diagonal matrix Λ which simultaneously diagonalize the two covariance matrices as

$$W'\mathbf{R_1}W = \Lambda_1 \quad \text{and} \quad W'\mathbf{R_2}W = I - \Lambda_1 = \Lambda_2 \tag{B.1}$$

where I is the identity matrix. This diagonalization is achieved whitening the composite matrix $\mathbf{R_s} = \mathbf{R_1} + \mathbf{R_2}$, i.e., obtaining a matrix P such that the covariance is equal to I

$$P'(\mathbf{R_1} + \mathbf{R_2})P = I. \tag{B.2}$$

The whitening matrix P is given by $P = U_s\sqrt{\Lambda_s^{-1}}$ where Λ_s is the diagonal matrix of eigenvalues and U_s is the matrix of eigenvectors of $\mathbf{R_s}$. Considering $S_1 = P'\mathbf{R_1}P$ and $S_2 = P'\overline{\mathbf{R}}_\mathbf{2}P$ then $S_1 + S_2 = I$. If S_1 is decomposed by

$$S_1 = \Phi\Lambda_1\Phi' \quad \text{or} \quad S_1\Phi = \Phi\Lambda_1 \tag{B.3}$$

where Φ and Λ_1 are respectively the eigenvectors and eigenvalue matrices of S_1, then

$$S_1\Phi + S_2\Phi = I\Phi \tag{B.4}$$

which leads to

$$\Lambda_2 = I - \Lambda_1 \tag{B.5}$$

where Λ_2 is the eigenvalue matrix of $S_2 = \Phi\Lambda_2\Phi'$. Therefore $\mathbf{R_1}$ and $\mathbf{R_2}$ share the same eigenvector matrix $W = P\Phi$, and the projection of data $\mathbf{X}$ is given according to

$$\mathbf{Y} = W'\mathbf{X}. \tag{B.6}$$

Assuming that the eigenvalues of $\mathbf{R_1}$ are in descending order $\lambda_1^{(1)} > \lambda_1^{(2)} > \cdots > \lambda_1^{(N)}$ then the eigenvalues of $\mathbf{R_2}$ will be reversely ordered. The eigenvector with largest eigenvalue

for one class has the smallest eigenvalue for the other class and vice versa. Each column vector (eigenvector) w_j $(j = 1, .., N)$ of the matrix W is a spatial filter. The first and last eigenvectors are the optimal eigenvectors that discriminate the two classes. The central eigenvectors maximize the common features of the two classes.

B.2 Rayleigh quotient and generalized eigenvalue problem

Consider the Rayleigh quotient

$$J(W) = \frac{W'\mathbf{R_1}W}{W'\mathbf{R_2}W}. \tag{B.7}$$

The optimal maximization of the Rayleigh quotient is given by differentiating $J(W)$ in order to W and equating to zero

$$\begin{aligned} \frac{\partial J(W)}{\partial W} =& \frac{\partial}{\partial W}\left(\frac{W'\mathbf{R_1}W}{W'\mathbf{R_2}W}\right) = 0 \\ \Leftrightarrow& (W'\mathbf{R_2}W)\frac{\partial}{\partial W}(W'\mathbf{R_1}W) - (W'\mathbf{R_2}W)\frac{\partial}{\partial W}(W'\mathbf{R_2}W) = 0 \\ \Leftrightarrow& (W'\mathbf{R_2}W)2\mathbf{R_1}W - (W'\mathbf{R_1}W)2\mathbf{R_2}W = 0 \end{aligned} \tag{B.8}$$

Dividing by $2W'\mathbf{R_2}W$, then

$$\begin{aligned} \mathbf{R_1}W - J(W)\mathbf{R_2}W &= 0 \\ \mathbf{R_2^{-1}R_1}W &= J(W)W \end{aligned} \tag{B.9}$$

Defining $J(w_i) \equiv \lambda_i$ then

$$\mathbf{R_2^{-1}R_1}w_i = \lambda_i w_i \tag{B.10}$$

which is a generalized eigenvalue problem.

B.3 Max-SNR beamformer and CSP

We show that the CSP solution is the same of that obtained from the optimization of the Rayleigh quotient. Let consider the Rayleigh quotient

$$J(W) = \frac{W'\mathbf{R_1}W}{W'\mathbf{R_s}W} \tag{B.11}$$

where $\mathbf{R_1}$ and $\mathbf{R_s} = \mathbf{R_1} + \mathbf{R_2}$ are the covariance matrices already defined in section B.1. According to section B.2, the optimal W is achieved by finding the generalized eigenvalue decomposition that satisfies

$$\mathbf{R_1}W = \mathbf{R_s}W\Lambda. \tag{B.12}$$

Substituting $\mathbf{R_s} = \mathbf{R_1} + \mathbf{R_2}$ in (B.12) we obtain

$$\begin{aligned} \mathbf{R_2} W \Lambda &= \mathbf{R_1} W (I - \Lambda) \\ \mathbf{R_2} W \Lambda (I - \Lambda)^{-1} &= \mathbf{R_1} W \end{aligned} \tag{B.13}$$

$$(\mathbf{R_2})^{-1} \mathbf{R_1} W = W \Lambda (I - \Lambda)^{-1} \tag{B.14}$$

and thus the eigenvectors W are obtained from the EVD of $(\mathbf{R_2})^{-1}\mathbf{R_1}$ or $(\mathbf{R_s})^{-1}\mathbf{R_1}$ which both share the same eigenvector matrix, provided that $\mathbf{R_2}$ and $\mathbf{R_s}$ are nonsingular. The projected variance ratio $v_t^{(k)}/v_{nt}^{(k)}$ is $\lambda^{(k)}/(1-\lambda^{(k)})$. If we impose $W'\mathbf{R_s}W = I$ in (B.12), then

$$\begin{aligned} \mathbf{R_1} W &= \mathbf{R_s} W \Lambda \\ W' \mathbf{R_1} W &= W' \mathbf{R_s} W \Lambda \\ W' \mathbf{R_1} W &= I \Lambda \\ \mathbf{R_1} &= W \Lambda W' \end{aligned} \tag{B.15}$$

and replacing $\mathbf{R_1} = W\Lambda W'$ in (B.13) we yield

$$\begin{aligned} \mathbf{R_2} W \Lambda &= W \Lambda W' W (I - \Lambda) \\ \mathbf{R_2} W &= W (I - \Lambda) \\ \mathbf{R_2} &= W (I - \Lambda) W' \end{aligned} \tag{B.16}$$

i.e., we reach the exact CSP solution (B.1) as we wanted to demonstrate.

The Max-SNR stated as the maximization of the Rayleigh quotient corresponds to a particular case of CSP, which considers only the larger eigenvalue that maximizes the discrimination power.

Appendix C

BCI-Tetris

This appendix succinctly describes a Tetris-like game controlled with BCI, developed in parallel with the main research work, and presents some qualitative results. This application highlights the BCI game research area, that we believe will have a significant growth in the next years.

C.1 *Case-study V* - Tetris game controlled with BCI

Is it possible to control a Tetris-like game with BCI? In order to answer this question, we adapted and simplified the original Tetris game to be controlled with BCI. Three different versions, offering different forms of control, were implemented as proof-of-concepts [Pires 2011a]. The graphical design of the proposed Tetris was developed by two undergraduate students under my co-supervision in their final MSc. thesis [Casaleiro 2010] and [Torres 2011]. The graphical design of these Tetris versions was developed in a TCL/Tk framework. Signal processing and classification algorithms were implemented in Simulink as in previous chapters. The data communication between the Simulink module and the TCL/Tk module was made through a shared memory driver, developed for this purpose. Two Tetris versions, denoted as V1 and V2, are based uniquely on P300 ERP elicited by an oddball paradigm, and a third version, denoted as V3, combines P300 with SMR induced by motor imagery.

C.1.1 P300-based Tetris: V1 and V2

In Tetris V1, there are 16 combinations representing different positions and rotations of four different pieces (Fig. C.1b)Left). The 16 pieces (combinations) flash randomly in pre-defined positions of the Tetris board. User has to select a target piece indicated at the bottom of the board (Fig. C.1a)). In Tetris V2, the selection of the piece is made in two instances. First, user selects the rotation of the piece and then, selects the horizontal position. At this phase of development, we had to explicitly provide the target piece at the bottom of the Tetris board, to clear up if the miss-selections came from miss-classifications of the algorithm or from wrong selections of the user. The pieces flash with a SOA of 200 ms. No significant differences exist between these paradigms and those described in previous chapters, except that symbols (pieces) are only visible at the time of flashing, and symbols are not associated with letters or words. Tetris V1 and V2 have different target probabilities, namely 1:16 and 1:4. The two versions were tested online by several participants after a calibration session (see a photo of an experimental session

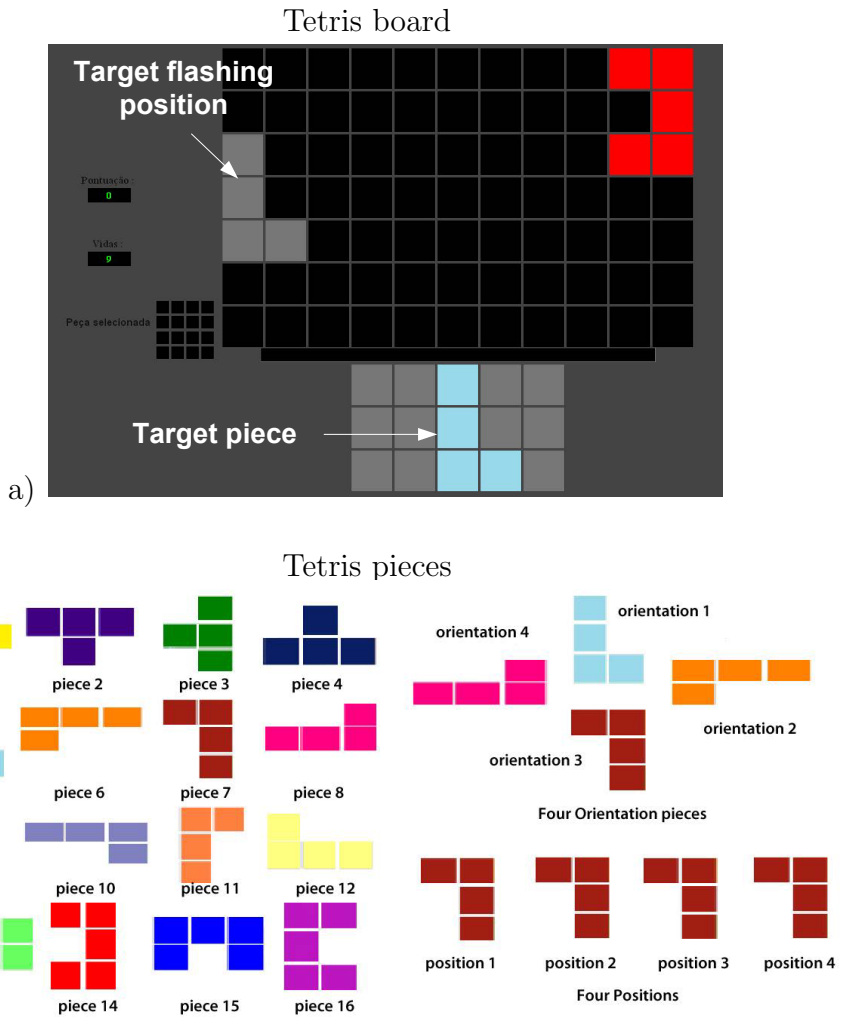
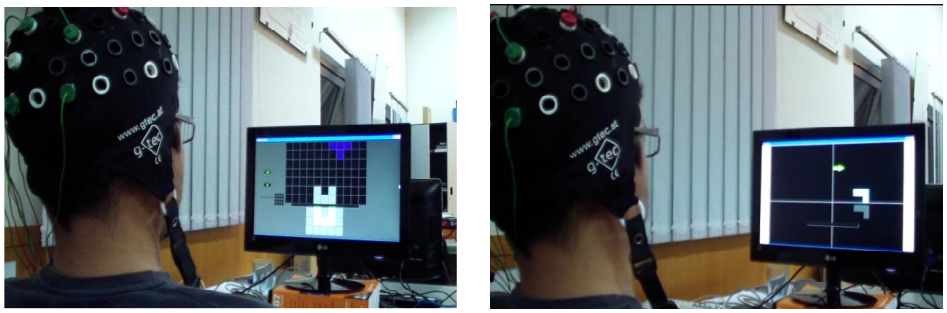


Figure C.1: a) Game board of the P300-based Tetris. The target piece to be selected is indicated at the bottom of the board. To help the player, the position where the piece is flashed appears in gray. The red piece is one of the pieces being flashed at the screenshot time; b) Left) 16 combinations of 4 pieces, combining simultaneously position and orientation; b) Right) Example of one piece showing the four possible positions along the horizontal axis and the four possible rotations.

in Fig.C.2a)). Both versions were effectively controlled. Nevertheless, participants performed worse with V2 than with V1, thus requiring a higher number of target repetitions. The higher target probability of version V2 may explain this difference. Additionally, in Tetris V2, participants failed more on the selection of horizontal positions than in the selection of the orientation. This may indicate that these two types of events have different perceptual effects.



a) b)

Figure C.2: Experimental sessions of user testing Tetris games. a) P300 Tetris (V1); P300-SMR Tetris (V3).

C.1.2 Combined P300-SMR: Tetris V3

P300-SMR Tetris is the most interesting, approaching the original Tetris game. The orientation of the piece is detected with the P300 neural mechanism, the same way as in Tetris V2, while the horizontal target position is reached by moving the piece to left or right, by performing motor imagery. Participants had to train how to control their SMR. The training session, of about 30 min, was made according to the widely used protocol described in [Pfurtscheller 1998]. During the training sessions, participants performed periods of motor imagery and periods of rest, according to the given cues (left, right and rest). Participants were asked not to move and to keep the arms and hands relaxed. At a left cue, participants had to imagine a left hand movement (e.g., opening and closing the left hand) and at right cue, they had to imagine a right hand movement (e.g., opening and closing the right hand). The type of motor imagination was only suggested, thus participants could choose their own imagination mechanisms. During rest periods, participants were asked to not imagine motor tasks. The training datasets consisted of 100 trials of left motor imagery and 100 trials of right motor imagery. Classification models, obtained from calibration, were fitted to the subject taking SMR features, namely the band power of μ and β rhythms [Pires 2011a].

After the training session, each participant played the game. After selecting the orientation of the piece, the user moved continuously the piece to left or right until the target position was reached (with visual neurofeedback). The displacement of the piece could be adjusted to be more or less reactive (taking more or less time), in order to fit the participant's performance. Successful results were achieved by several participants, although some couldn't learn how to control their rhythms (see a photo of an experimental session in Fig.C.2b)).

C.1.3 Discussion

The idea behind the implementation of the BCI-Tetris was to develop a game that could be used clinically by children or adolescents with ADHD. Neurofeedback, the real-time feedback of brain activity, is currently being used in many clinics for the treatment of

ADHD, as an alternative approach to pharmacologic treatments [Monastra 2005]. Neurofeedback is provided in the form of game control to increase the motivation of the trainees [Pope 2004]. Rhythms associated with attention are translated to a command that controls a specific variable of the game, providing visual feedback of neural activity, thereby inducing the user to modulate his rhythms. Our proposed Tetris games are substantially different concerning the control form and the level of difficulty. For the best of our knowledge, in the context of ADHD, P300 oddball paradigms have never been used for the treatment of ADHD. P300 has been used only as a neurophysiologic marker of ADHD or to assess the improvements of patients after the treatment. Since P300 is an endogenous ERP that depends on selective attention, we can hypothesize that continuous sessions playing P300-based games could improve the attention levels of the players. This supposition obviously still lacks of validation by neurotherapists. Moreover, motor imagery is also not a common practice used in ADHD. An effective clinical application of this game would obviously require several modifications and adjustments based on neurotherapists experience.

As low-cost commercial EEG headsets appear in the market, such as Emotiv [Emotiv 2011] and Neurosky [Neurosky 2011], simple BCI game applications are emerging for entertainment. BCI can provide a new control dimension, freeing the hands of the players to control some inputs of the game, while BCI is used to control others inputs.

Appendix D

Automatic detection of P300 evoked by subliminar visual tasks

This appendix succinctly describes a case-study that investigates the relation between psychophysical detection of visual salient features and their neurophysiological correlates at the event level. Our proposed signal processing and classification algorithms are applied for automatic detection of P300 ERPs in neurophysiologic responses, thereby validating the link between psychophysical and neurophysiological responses. Just a few of the overall results are here presented.

D.1 *Case-Study VI* - Automatic detection of P300 evoked by subliminar visual tasks

An experimental task was designed by Marta Teixeira [Teixeira 2010] [Teixeira 2011] at IBILI, assessing psychophysical and neurophysiological aspects related to perceptual saliency of visual targets. The experimental task investigates whether the P300 signal is modulated by the degree of perceptual saliency, independently of visual target properties. To confirm if there is a direct link between psychophysics and the neurophysiology, we applied our signal processing and classification algorithms to detect P300 ERPs in neurophysiologic responses. My contribution to this work was on the implementation of the signal processing and classification algorithms [Teixeira 2011]. Classification was applied at a single-trial level, as a psychophysiological measure of perception. If the algorithms can efficiently detect P300, this experiment might be used as a clinical test for an objective visual acuity measurement, independently of the cooperation of the patient.

D.1.1 Task description

Two experiments, both consisting of an oddball design, were tested by 10 healthy participants. Experiment 1 was designed to test the effect of perceptual saliency of color/luminance and texture on P300 amplitude and latency. Saliency was manipulated by having 3 target levels of perceptual deviation from a standard red coloured patch (see Fig. D.1a)). Texture was manipulated by using 3 target levels of phase offset deviation against a standard with no phase offset (see Fig. D.1b)). In experiment 2, saliency was manipulated by having 3 target levels of isoluminant chromatic deviation against a green squared patch, for the chrominance condition (see Fig. D.2a)). Luminance was manipulated by using 3 target levels of luminance deviant patches (see Fig. D.2b)). Targets are

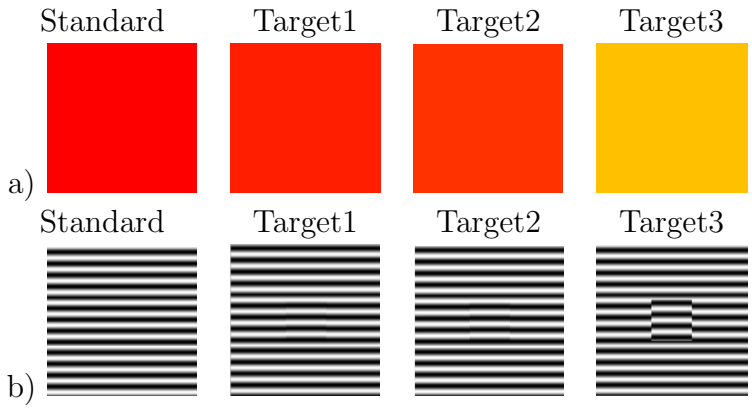


Figure D.1: Experiment 1: a) Stimuli for colour/luminance saliency; b) Stimuli for texture saliency. Images taken from [Teixeira 2011].

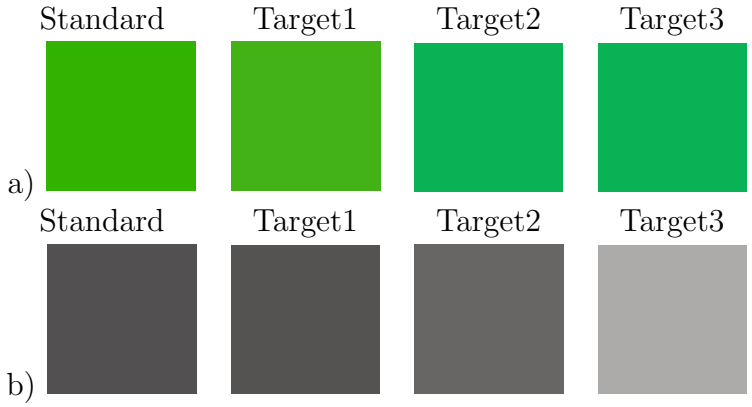


Figure D.2: Experiment 2: a) Stimuli for chrominance saliency; b) Stimuli for luminance saliency. Images taken from [Teixeira 2011].

designated as Target1, Target2, and Target3 according to their increased level of salience. Each deviant target had a probability of 1:16, with a minimal inter-target distance of 6 stimulus repetitions. Stimuli were displayed for 500 ms, with an ISI of 1500 ms. During the experiments, participants had to respond, by clicking a button, for each detected deviant target.

D.1.2 P300 classification

Detection of P300 was performed on a single-trial basis, i.e., taking a single repetition of each target stimulus. The classification process was based on the proposed C-FMS spatial filter and on the LDA classifier. Two classification scenarios were analysed. In the first one, classification models were fitted individually to each participant. The main goal was to infer if the automatic P300 detection matched the behavioural response (inferred from the motor response of participant). As an example, Fig. D.3 compares the single trial detection with motor response for the chrominance experiment. Similar results were

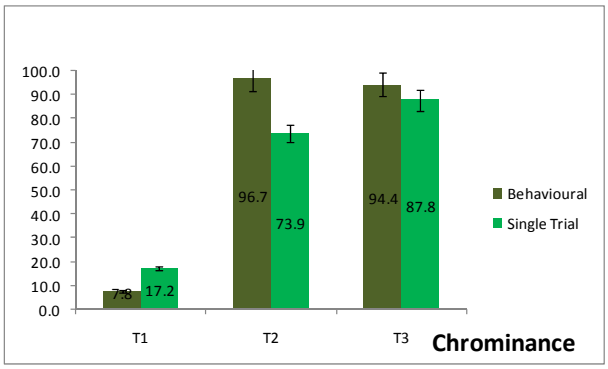


Figure D.3: Detection rate of behavioural vs. single trial in the chrominance experiment. Graph taken from [Teixeira 2011].

Table D.1: Classification results (balanced accuracy) of independent model for the chrominance condition.

	Subj 1	Subj 2	Subj 3	Subj 4	Subj 5	Subj 6	Subj 7	Subj 8	Subj 9	Subj 10	average
Target 3	88.77	94.79	90.05	62.66	84.14	85.26	77.80	87.45	96.70	89.41	82.40
Target 2	85.89	89.57	73.18	54.63	76.74	73.28	72.80	84.26	85.47	74.40	72.76
Target 1	75.41	81.06	45.10	47.77	44.99	47.38	53.02	44.26	65.20	45.20	46.82

achieved for the other conditions. Classification results were obtained individually for each subject through leave-one-out (LOO) cross-validation applied to all epochs of the datasets. Statistical tests confirmed the correlation between single-trial detection vs behavioural response. In the second classification scenario, results were obtained from an independent model built from a set of three subjects and tested on the remaining subjects. The model was obtained through LOO cross-validation within the set of participants, by choosing the three subjects that best generalized to the other subjects. The model was created joining the datasets of these three subjects. This classification scenario validates the algorithms performance on unseen subject's data. This scenario is obviously the most important for clinical application, as an independent model is required for a vision test. Results are shown in Table D.1 for the chrominance experiment. The average results were obtained discarding the three subjects from which the model was obtained (subjects 1, 2 and 9). The results are only slightly worse than in the first scenario, showing that the independent model generalizes well to other subjects. Similar results were obtained for the other experiments. This infers the good generalization properties of the C-FMS spatial filter approach, and validates its potential use for a clinical acuity vision test.

References

[Ahi 2011] S.T. Ahi, H. Kambara and Y. Koike. *A Dictionary-Driven P300 Speller With a Modified Interface*. IEEE Trans. Neural Syst. Rehab. Eng., vol. 19, no. 1, pages 6 –14, feb. 2011.

[Allison 2003a] B Z Allison. *P3 or not P3: toward a better P300 BCI*. Phd dissertation, University of California, 2003.

[Allison 2003b] B.Z. Allison and J.A. Pineda. *ERPs evoked by different matrix sizes: implications for a brain computer interface (BCI) system*. IEEE Trans. Neural Syst. Rehab. Eng., vol. 11, no. 2, pages 110 –113, june 2003.

[Allison 2006] Brendan Z. Allison and Jaime A. Pineda. *Effects of SOA and flash pattern manipulations on ERPs, performance, and preference: Implications for a BCI system*. Int. J. Psychophysiology, vol. 59, no. 2, pages 127 – 140, 2006.

[Allison 2008] Brendan Z. Allison, Dennis J. McFarland, Gerwin Schalk, Shi Dong Zheng, Melody Moore Jackson and Jonathan R. Wolpaw. *Towards an independent brain-computer interface using steady state visual evoked potentials*. Clinical Neurophysiology, vol. 119, no. 2, pages 399 – 408, 2008.

[Bashashati 2007] Ali Bashashati, Mehrdad Fatourechi, Rabab K Ward and Gary E Birch. *A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals*. J. Neural Eng., vol. 4, no. 2, page R32, 2007.

[Bax 2005] M Bax, M. Goldstein, P. Rosenbaum, A. Leviton, N. Paneth, B Dan, B Jacobsson and D Damiano. *Proposed definition and classification of cerebral palsy, April 2005*. Developmental Medicine Child Neurology, vol. 47, no. 8, pages 571–571, 2005.

[Bayliss 2003] J.D. Bayliss. *Use of the evoked potential P3 component for control in a virtual apartment*. IEEE Trans. Neural Syst. Rehab. Eng., vol. 11, no. 2, pages 113 –116, june 2003.

[BCI-Competition 2003] BCI-Competition. *BCI Competition 2003 - Data set IIb*. `http://www.bbci.de/competition/ii/`, 2003.

[Bendat 2000] J. S. Bendat and A. G. Piersol. Random data - analysis and measurement procedure. John Wiley & Sons, Inc, 2000.

[Birbaumer 2000] N. Birbaumer, A. Kubler, N. Ghanayim, T. Hinterberger, J. Perelmouter, J. Kaiser, I. Iversen, B. Kotchoubey, N. Neumann and H. Flor. *The thought translation device (TTD) for completely paralyzed patients*. IEEE Trans. Rehab. Eng., vol. 8, no. 2, pages 190–193, 2000.

[Birbaumer 2006] Niels Birbaumer. *Breaking the silence: Brain-computer interfaces (BCI) for communication and motor control.* Psychophysiology, vol. 43, no. 6, pages 517–532, 2006.

[Blanchard 2004] G. Blanchard and B. Blankertz. *BCI competition 2003-data set IIa: spatial patterns of self-controlled brain rhythm modulations.* IEEE Trans. Biomed. Eng., vol. 51, no. 6, pages 1062–1066, June 2004.

[Blankertz 2006] B. Blankertz, G. Dornhege, M. Krauledat, K.-R. Müller, V. Kunzmann, F. Losch and G. Curio. *The Berlin brain-computer interface: EEG-based communication without subject training.* IEEE Trans. Neural Syst. Rehab. Eng., vol. 14, no. 2, pages 147 –152, june 2006.

[Blankertz 2010] Benjamin Blankertz, Michael Tangermann, Carmen Vidaurre, Siamac Fazli, Claudia Sannelli, Stefan Haufe, Cecilia Maeder, Lenny E Ramsey, Irene Sturm, Gabriel Curio and Klaus R Müeller. *The Berlin Brain-Computer Interface: Non-Medical Uses of BCI Technology.* Frontiers Neurosc., vol. 4, 2010.

[Blankertz 2011] Benjamin Blankertz, Steven Lemm, Matthias Treder, Stefan Haufe and Klaus-Robert Müller. *Single-trial analysis and classification of ERP components - A tutorial.* NeuroImage, vol. 56, no. 2, pages 814 – 825, 2011.

[BMBF 2005] BMBF. *Educational Research and Neurosciences - Expectations, Evidence, Research Prospects.* Technical report, German Bundesministerium - Federal Ministry of Education and Research (BMBF), 2005.

[Boser 1992] Bernhard E. Boser, Isabelle M. Guyon and Vladimir N. Vapnik. *A Training Algorithm for Optimal Margin Classifiers.* In David Haussler, editor, Proc. 5th Annual Workshop on Computational Learning Theory (COLT'92), pages 144–152. ACM Press, New York, NY, USA, 1992.

[BrainAble 2010] Consortium BrainAble. *BCI and BNCI Systems, State of the Art.* Technical report, Graz University of Technology, 2010.

[Brunner 2010] P Brunner, S Joshi, S Briskin, J R Wolpaw, H Bischof and G Schalk. *Does the P300 speller depend on eye gaze?* J. Neural Eng., vol. 7, no. 5, 2010.

[Burges 1998] Christopher J. C. Burges. *A Tutorial on Support Vector Machines for Pattern Recognition.* Data Min. Knowl. Discov., vol. 2, no. 2, pages 121–167, 1998.

[Carmena 2003] Jose M Carmena, Mikhail A Lebedev, Roy E Crist, Joseph E O'Doherty, David M Santucci, Dragan F Dimitrov, Parag G Patil, Craig S Henriquez and Miguel A. L Nicolelis. *Learning to Control a Brain-Machine Interface for Reaching and Grasping by Primates.* PLoS Biol, vol. 1, no. 2, 10 2003.

[Casaleiro 2010] Nuno Casaleiro. Desenvolvimento de paradigmas visuais para aplicacoes de interface cerebro-computador baseadas em sinais electroencefalograficos. Msc dissertation, Universidade de Coimbra - Departamento de Engenharia Electrotecnica, 2010.

[Cecotti 2010] H. Cecotti. *A Self-Paced and Calibration-Less SSVEP-Based Brain-Computer Interface Speller*. IEEE Trans. Neural Syst. Rehab. Eng., vol. 18, no. 2, pages 127 –133, april 2010.

[Cecotti 2011] H. Cecotti and A. Graser. *Convolutional Neural Networks for P300 Detection with Application to Brain-Computer Interfaces*. IEEE Trans. Pattern Analysis Machine Intell., vol. 33, no. 3, pages 433 –445, march 2011.

[Cedarbaum 1999] Jesse M. Cedarbaum, Nancy Stambler, Errol Malta, Cynthia Fuller, Dana Hilt, Barbara Thurmond and et al. *The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function*. J. Neurolog. Sciences, vol. 169, no. 1-2, pages 13 – 21, 1999.

[Cincotti 2008] Febo Cincotti, Donatella Mattia, Fabio Aloise, Simona Bufalari, Gerwin Schalk, Giuseppe Oriolo, Andrea Cherubini, Maria Grazia Marciani and Fabio Babiloni. *Non-invasive brain-computer interface system: Towards its application as assistive technology*. Brain Research Bulletin, vol. 75, no. 6, pages 796 – 803, 2008.

[Comerchero 1999] Marco D. Comerchero and John Polich. *P3a and P3b from typical auditory and visual stimuli*. Clinical Neurophysiology, vol. 110, no. 1, pages 24–30, January 1999.

[Correa 2006] Angel Correa, Juan Lupianez, Eduardo Madrid and Pio Tudela. *Temporal attention enhances early visual processing: A review and new evidence from event-related potentials*. Brain Research, vol. 1076, no. 1, pages 116 – 128, 2006.

[Coyle 2007] Shirley M Coyle, Tomas E Ward and Charles M Markham. *Brain-computer interface using a simplified functional near-infrared spectroscopy system*. J. Neural Eng., vol. 4, no. 3, page 219, 2007.

[Cui 1999] R. Q. Cui, D. Huter and L. Deecke. *Neuroimage of Voluntary Movement: Topography of the Bereitschaftspotential, a 64-Channel DC current source study*. Neuroimage, vol. 9, pages 124–134, 1999.

[Cyrulnik 2008] Shana E. Cyrulnik and Veronica J. Hinton. *Duchenne muscular dystrophy: A cerebellar disorder?* Neurosc. Biobehavioral Reviews, vol. 32, no. 3, pages 486 – 496, 2008.

[Dal Seno 2010] B. Dal Seno, M. Matteucci and L.T. Mainardi. *The Utility Metric: A Novel Method to Assess the Overall Performance of Discrete Brain-Computer Interfaces*. IEEE Trans. Neural Syst. Rehab. Eng., vol. 18, no. 1, pages 20–28, feb. 2010.

[Daly 2006] J.J. Daly, Yin Fang, E.M. Perepezko, V. Siemionow and G.H. Yue. *Prolonged cognitive planning time, elevated cognitive effort, and relationship to coordination and motor control following stroke.* IEEE Trans. Neural Syst. Rehab. Eng., vol. 14, no. 2, pages 168 –171, june 2006.

[Daly 2008] Janis J Daly and Jonathan R Wolpaw. *Brain-computer interfaces in neurological rehabilitation.* The Lancet Neurology, vol. 7, no. 11, pages 1032–1043, 2008.

[Daly 2009] Janis Daly, Roger Cheng, Jean Rogers, Krisanne Litinas, Kenneth Hrovat and Mark Dohring. *Feasibility of a New Application of Noninvasive Brain Computer Interface (BCI): A Case Study of Training for Recovery of Volitional Motor Control After Stroke.* J. Neurologic Physical Therapy, vol. 33, no. 4, pages 203–211, 2009.

[de Cheveigne 2008] Alain de Cheveigne and Jonathan Z. Simon. *Denoising based on spatial filtering.* J. Neurosc. Methods, vol. 171, no. 2, pages 331–339, 2008.

[Ding 2005] Chris Ding and Hanchuan Peng. *Minimum Redundancy Feature Selection from Microarray Gene Expression Data.* J. Bioinform. Comput. Biol., vol. 3, no. 2, pages 185–205, 2005.

[Donchin 2000] E. Donchin, K. Spencer and Wijesinghe R. *The Mental Prosthesis: Assessing the Speed of a P300-Based Brain-Computer Interface.* IEEE Trans. Rehabil. Eng., vol. 8, no. 2, pages 174–179, June 2000.

[Dornhege 2006] G. Dornhege, B. Blankertz, M. Krauledat, F. Losch, G. Curio and K.-R. Müller. *Combined Optimization of Spatial and Temporal Filters for Improving Brain-Computer Interfacing.* IEEE Trans. Biomed. Eng., vol. 53, no. 11, pages 2274–2281, Nov. 2006.

[Duda 2001] R. Duda, P. Hart and D. Stork. Pattern classification. Wiley London UK, 2001.

[Emotiv 2011] Emotiv. *Emotiv Epoc neuroheadset.* `http://www.emotiv.com`, 2011. Accessed in 2011.

[Fabiani 2004] G. Fabiani, D. J. McFarland, J. R. Wolpaw and P. Pfurtscheller. *Conversion of EEG Activity Into Cursor Movement by Brain-Computer Interface (BCI).* IEEE Trans. Neural Syst. Rehab. Eng., vol. 12, no. 3, pages 331–338, September 2004.

[Farwell 1988] L.A Farwell and E. Donchin. *Talking off the top of your head: toward a mental prosthesis utilizing event related brain potentials.* Electroencephalography Clinical Neurophysiology, vol. 70, no. 6, pages 510–523, 1988.

[Finke 2009] Andrea Finke, Alexander Lenhardt and Helge Ritter. *The MindGame: A P300-based brain-computer interface game*. Neural Networks, vol. 22, no. 9, pages 1329 – 1333, 2009.

[Fiorentino 2009] G. Fiorentino, P. Carideo, A. Annunziata, M. Boselli, R. Cauteruccio and M. Caputi. *s dyspnea in amyotrophic lateral sclerosis patients a signal for bad nutrition?* Nutritional Therapy Metabolism, vol. 27, no. 2, pages 100–104, 2009.

[Frenzel 2010] Stefan Frenzel and Elke Neubert. *Is the P300 Speller Independent?* Computing Research Repository (CoRR), 2010.

[Friedman 2001] David Friedman, Yael M. Cycowicz and Helen Gaeta. *The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty*. Neurosc. Biobehavioral Reviews, vol. 24, no. 4, pages 355–373, 2001.

[Fukunaga 1970] K. Fukunaga and W. G. Koontz. *Application of the Karhunen-Loeve Expansion to Feature Selection and Ordering*. IEEE Trans. Comput., vol. C-19, no. 4, pages 311–318, April 1970.

[Fukunaga 1990] K. Fukunaga. Introduction to statistical pattern recognition, second edition. Morgan Kaufmann - Academic Press, 1990.

[Furdea 2009] A. Furdea, S. Halder, D.J. Krusienski, D. Bross, F. Nijboer, N. Birbaumer and A. Kübler. *An auditory oddball (P300) spelling system for brain-computer interfaces*. Psychophysiology, vol. 46, no. 3, pages 617–625, 2009.

[Galán 2008] F. Galán, M. Nuttin, E. Lew, P.W. Ferrez, G. Vanacker, J. Philips and J. del R. Millán. *A brain-actuated wheelchair: Asynchronous and non-invasive Brain-computer interfaces for continuous control of robots*. Clinical Neurophysiology, vol. 119, no. 9, pages 2159 – 2169, 2008.

[Gao 2003] X. Gao, D. Xu, M. Cheng and S. Gao. *A BCI-based environmental controller for the motion disabled*. IEEE Trans. Neural Syst. Rehab. Eng., vol. 11, no. 2, pages 137–140, June 2003.

[Gevensleben 2009] Holger Gevensleben, Birgit Holl, Björn Albrecht, Claudia Vogel, Dieter Schlamp, Oliver Kratz, Petra Studer, Aribert Rothenberger, Gunther H. Moll and Hartmut Heinrich. *Is neurofeedback an efficacious treatment for ADHD? A randomised controlled clinical trial*. J. Child Psychology and Psychiatry, vol. 50, no. 7, pages 780–789, 2009.

[Grosse-Wentrup 2009] M. Grosse-Wentrup, C. Liefhold, K. Gramann and M. Buss. *Beamforming in Noninvasive Brain-Computer Interfaces*. IEEE Trans. Biomed. Eng., vol. 56, no. 4, pages 1209 –1219, april 2009.

[g.tec] g.tec. *gUSBamp amplifier*. `http://www.gtec.at`.

[gtec 2009] gtec. *Intendix - Personnal EEG-based Spelling System.* `http://www.intendix.com/`, 2009.

[Guan 2004] Cuntai Guan, M. Thulasidas and Jiankang Wu. *High performance P300 speller for brain-computer interface.* In IEEE Int. Workshop Biomed. Circuits Syst., pages 13–16, 2004.

[Guger 2009] Christoph Guger, Shahab Daban, Eric Sellers, Clemens Holzner, Gunther Krausz, Roberta Carabalona and et al. *How many people are able to control a P300-based brain-computer interface (BCI)?* Neurosci. Letters, vol. 462, no. 1, pages 94 – 98, 2009.

[Guo 2010] Jing Guo, Shangkai Gao and Bo Hong. *An Auditory Brain-Computer Interface Using Active Mental Response.* IEEE Trans. Neural Syst. Rehab. Eng., vol. 18, no. 3, pages 230 –235, june 2010.

[Guyon 2003] Isabelle Guyon and Andre Elisseeff. *An introduction to variable and feature selection.* J. Mach. Learn. Res., vol. 3, pages 1157–1182, 2003.

[Hall 2000] Mark A. Hall. *Correlation-based Feature Selection for Discrete and Numeric Machine Learning.* Proc. 17th Int. Conf. Machine Learning, pages 359–366, 2000.

[Haykin 1996] Simon Haykin. Adaptive filter theory (3rd ed.). Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1996.

[Heinrich 2008] Sven Heinrich and Michael Bach. *Signal and noise in P300 recordings to visual stimuli.* Doc. Ophthalmology, vol. 117, pages 73–83, 2008.

[Hinterberger 2004] T. Hinterberger, S. Schmidt, N. Neumann, J. Mellinger, B. Blankertz, G. Curio and N. Birbaumer. *Brain-computer communication and slow cortical potentials.* IEEE Trans. Biomed. Eng., vol. 51, no. 6, pages 1011 –1018, june 2004.

[Hoffmann 2006] U. Hoffmann, J. Vesin and T. Ebrahimi. *Spatial filters for the classification of event-related potentials.* In European Symposium on Artificial Neural Networks, 2006.

[Hoffmann 2008a] U. Hoffmann, A. Yazdani, J M. Vesin and T Ebrahimi. *Bayesian Feature Selection Applied In a P300 Brain-Computer Interface.* In 16th European Signal Process. Conf., 2008.

[Hoffmann 2008b] Ulrich Hoffmann, Jean-Marc Vesin, Touradj Ebrahimi and Karin Diserens. *An efficient P300-based brain-computer interface for disabled subjects.* J Neuroscience Methods, vol. 167, no. 1, pages 115–125, 2008.

[Hohne 2011] Johannes Hohne, Martijn Schreuder, Benjamin Blankertz and Michael Tangermann. *A novel 9-class auditory ERP paradigm driving a predictive text entry system.* Frontiers Neurosc., vol. 5, no. 0, 2011.

[Ioannis 1998] Milonas Ioannis. *Amyotrophic lateral sclerosis: an introduction.* J. Neurology, vol. 245, pages S1–S3, 1998.

[Iturrate 2009] I. Iturrate, J.M. Antelis, A. Kubler and J. Minguez. *A Noninvasive Brain-Actuated Wheelchair Based on a P300 Neurophysiological Protocol and Automated Navigation.* IEEE Trans. Robotics, vol. 25, no. 3, pages 614–627, June 2009.

[Ivannikov 2009] Andriy Ivannikov, Igor Kalyakin, Jarmo Hamalainen, Paavo H.T. Leppanen, Tapani Ristaniem, Heikki Lyytinen and Tommi Karkkainen. *ERP denoising in multichannel EEG data using contrasts between signal and noise subspaces.* J Neurosc. Methods, vol. 180, no. 2, pages 340–351, 2009.

[Jin 2011a] Jing Jin, Brendan Allison, Eric Sellers, Clemens Brunner, Petar Horki, Xingyu Wang and et al. *Optimized stimulus presentation patterns for an event-related potential EEG-based brain-computer interface.* Med. Biol. Eng. Comput., vol. 49, pages 181–91, 2011.

[Jin 2011b] Jing Jin, Brendan Z Allison, Eric W Sellers, Clemens Brunner, Petar Horki, Xingyu Wang and Christa Neuper. *An adaptive P300-based control system.* J. Neural Eng., vol. 8, no. 3, page 036006, 2011.

[Jones 2007] Martha Wilson Jones, Elaine Morgan, Jean E. Shelton and Christine Thorogood. *Cerebral Palsy: Introduction and Diagnosis (Part I).* J. Pediatric Health Care, vol. 21, no. 3, pages 146 – 152, 2007.

[Jung 1979] Richard Jung and Wiltrud Berger. *Hans Bergers Entdeckung des Elektrenkephalogramms und seine ersten Befunde 1924-1931.* European Archives Psychiatry Clinical Neurosc., vol. 227, pages 279–300, 1979.

[Jung 2000] Tzyy-Ping Jung, Scott Makeig, Marissa Westerfield, Jeanne Townsend, Eric Courchesne and Terrence J. Sejnowski. *Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects.* Clinical Neurophysiology, vol. 111, no. 10, pages 1745–1758, October 2000.

[Kaper 2004] M. Kaper, P. Meinicke, U. Grossekathoefer, T. Lingner and H. Ritter. *BCI competition 2003-data set IIb: support vector machines for the P300 speller paradigm.* IEEE Trans. Biomed Eng., vol. 51, no. 6, pages 1073–1076, June 2004.

[Kelly 2005] S.P. Kelly, E.C. Lalor, C. Finucane, G. McDarby and R.B. Reilly. *Visual spatial attention control in an independent brain-computer interface.* IEEE Trans. Biomed. Eng., vol. 52, no. 9, pages 1588 –1596, sept. 2005.

[Kim 2008] Sung-Phil Kim, John D Simeral, Leigh R Hochberg and John P Donoghueand Michael J Black. *Neural control of computer cursor velocity by decoding motor cortical spiking activity in humans with tetraplegia.* J. Neural Eng., no. 5, pages 455–476, 2008.

[Kohlmorgen 2004] J. Kohlmorgen and B. Blankertz. *Bayesian classification of single-trial event-related potentials in EEG.* Int. J. Bif. Chaos, vol. 14, no. 2, pages 719–726, 2004.

[Krauledat 2008] Matthias Krauledat, Michael Tangermann, Benjamin Blankertz and Klaus-Robert Müller. *Towards Zero Training for Brain-Computer Interfacing.* PLoS ONE, vol. 3, no. 8, page e2967, 08 2008.

[Krigger 2006] K. Krigger. *Cerebral Palsy: An Overview.* American Family Physician, vol. 73, no. 1, pages 91–100, 2006.

[Krusienski 2006] D. Krusienski, E. Sellers, F. Cabestaing, S. Bayoudh, Dennis McFarland, Theresa Vaughan and Jonathan Wolpaw. *A comparison of classification techniques for the P300 Speller.* J. Neural Eng., no. 3, pages 299–305, 2006.

[Krusienski 2007] D. J. Krusienski, W. Sellers and T. M. Vaughan. *Common Spatio-Temporal Patterns for the P300 Speller.* 3rd Int. IEEE Conf. Neural Eng., NER07, pages 421–424, May 2007.

[Krusienski 2008] D. Krusienski, S. Sellers and J. Wolpaw D. McFarland T. Vaughan. *Toward Enhanced P300 Speller Performance.* J. Neurosc. Methods, vol. 167, no. 1, pages 15–21, January 2008.

[Krusienski 2011] Dean J Krusienski, Moritz Grosse-Wentrup, Ferran Galán, Damien Coyle, Kai J Miller, Elliott Forney and Charles W Anderson. *Critical issues in state-of-the-art brain-computer interface signal processing.* J. Neural Eng., vol. 8, no. 2, page 025002, 2011.

[Kübler 2001] Andrea Kübler, Nicola Neumann, Jochen Kaiser, Boris Kotchoubey, Thilo Hinterberger and Niels P. Birbaumer. *Brain-computer communication: Self-regulation of slow cortical potentials for verbal communication.* Archives Physical Medicine Rehab., vol. 82, no. 11, pages 1533 – 1539, 2001.

[Kubler 2008] A. Kubler and N. Birbaumer. *Brain-computer interfaces and communication in paralysis: Extinction of goal directed thinking in completely paralysed patients?* Clinical Neurophysiology, vol. 119, no. 11, pages 2658 – 2666, 2008.

[Leigh 1994] P. N. Leigh and K. Ray-Chaudhuri. *Motor neuron disease.* J. Neurology, Neurosurgery, Psychiatry, vol. 57, no. 8, pages 886–896, August 1994.

[Lemm 2005] S. Lemm, B. Blankertz, G. Curio and K.-R. Müller. *Spatio-spectral filters for improving the classification of single trial EEG.* IEEE Trans. Biomed. Eng., vol. 52, no. 9, pages 1541–1548, Sept. 2005.

[Lemm 2006] S. Lemm, G. Curio, Y. Hlushchuk and K.-R. Müller. *Enhancing the signal-to-noise ratio of ICA-based extracted ERPs.* IEEE Trans. Biomed Eng., vol. 53, no. 4, pages 601–607, April 2006.

[Lenhardt 2008] A. Lenhardt, M. Kaper and H.J. Ritter. *An Adaptive P300-Based Online Brain-Computer Interface*. IEEE Trans. Neural Syst. Rehab. Eng., vol. 16, no. 2, pages 121–130, April 2008.

[Leuthardt 2006] E.C. Leuthardt, K.J. Miller, G. Schalk, R.P.N. Rao and J.G. Ojemann. *Electrocorticography-based brain computer Interface-the seattle experience*. IEEE Trans. Neural Syst. Rehab. Eng., vol. 14, no. 2, pages 194 –198, june 2006.

[Li 2004] Yong Li, Xiaorong Gao, Hesheng Liu and Shangkai Gao. *Classification of single-trial electroencephalogram during finger movement*. IEEE Trans. Biomed. Eng., vol. 51, no. 6, pages 1019–1025, June 2004.

[Liu 2005] Huan Liu and Lei Yu. *Toward integrating feature selection algorithms for classification and clustering*. IEEE Trans. Knowledge Data Eng., vol. 17, no. 4, pages 491–502, April 2005.

[Liu 2010] Tao Liu, Leslie Goldberg, Shangkai Gao and Bo Hong. *An online brain-computer interface using non-flashing visual evoked potentials*. J. Neural Eng., vol. 7, no. 3, page 036003, 2010.

[Liu 2011] Yang Liu, Zongtan Zhou and Dewen Hu. *Gaze independent brain-computer speller with covert visual search tasks*. Clinical Neurophysiology, vol. 122, pages 1127–1136, 2011.

[Lopes 2011] Ana Lopes, Gabriel Pires, Luis Vaz and Urbano Nunes. *Wheelchair Navigation Assisted by Human-machine Shared-control and a P300-based BCI*. IEEE/RSJ Int Conf Intell Robots Syst IROS2011, 2011.

[Lotte 2007] F Lotte, M Congedo, A Lecuyer, F Lamarche and B Arnaldi. *A review of classification algorithms for EEG-based brain-computer interfaces*. J. Neural Eng., vol. 4, no. 2, page R1, 2007.

[Lu 2009] Shijian Lu, Cuntai Guan and Haihong Zhang. *Unsupervised Brain Computer Interface Based on Intersubject Information and Online Adaptation*. IEEE Trans. Neural Syst. Rehab. Eng., vol. 17, no. 2, pages 135 –145, april 2009.

[Luck 1994] Steven J. Luck and Steven A. Hillyard. *Spatial Filtering During Visual Search: Evidence From Human Electrophysiology*. J. Experimental Psychol.: Human Perception and Performance, vol. 20, no. 5, pages 1000 – 1014, 1994.

[Mak 2009] J.N. Mak and J.R. Wolpaw. *Clinical Applications of Brain-Computer Interfaces: Current State and Future Prospects*. IEEE Rev. Biomed. Eng., vol. 2, pages 187 –199, 2009.

[Makeig 1999] Scott Makeig, Marissa Westerfield, Jeanne Townsend, Tzyy-Ping Jung, Eric Courchesne and Terrence J. Sejnowski. *Functionally Independent Components of Early Event-Related Potentials in a Visual Spatial Attention Task*. Philos. Trans. Royal Society: Biol. Sciences, vol. 354, no. 1387, pages 1135–1144, 1999.

[Martens 2009] S M M Martens, N J Hill, J Farquhar and B Scholkopf. *Overlap and refractory effects in a brain-computer interface speller based on the visual P300 event-related potential.* J. Neural Eng., vol. 6, 2009.

[Martens 2010] S M M Martens and J M Leiva. *A generative model approach for decoding in the visual event-related potential-based brain-computer interface speller.* J. Neural Eng., vol. 7, no. 2, page 026003, 2010.

[Mason 2000] S.G. Mason and G.E. Birch. *A brain-controlled switch for asynchronous control applications.* IEEE Trans. Biomed. Eng., vol. 47, no. 10, pages 1297 –1307, oct 2000.

[Mathworks 2010a] Mathworks. Matlab version 7.10.0 (r2010a). The MathWorks Inc., Natick, Massachusetts, 2010.

[Mathworks 2010b] Mathworks. *Writing S-functions, Simulink - Dynamic System Simulation for Matlab*, 2010.

[McFarland 1997] D. J. McFarland, L.M. McCane, S. V. David and J. R. Wolpaw. *Spatial filter selection for EEG-based communication.* Electroencephalography Clinical Neurophysiology, vol. 103, pages 386–394, 1997.

[McFarland 2008a] Dennis J McFarland, Dean J Krusienski, William A Sarnacki and Jonathan R Wolpaw. *Emulation of computer mouse control with a noninvasive brain-computer interface.* J. Neural Eng., vol. 5, no. 2, 2008.

[McFarland 2008b] D.J. McFarland and J.R. Wolpaw. *Brain-Computer Interface Operation of Robotic and Prosthetic Devices.* IEEE Computer, vol. 41, no. 10, pages 52 –56, oct. 2008.

[McFarland 2010] Dennis J McFarland, William A Sarnacki and Jonathan R Wolpaw. *Electroencephalographic (EEG) control of three-dimensional movement.* J. Neural Eng., vol. 7, no. 3, 2010.

[Meinicke 2002] Peter Meinicke, Matthias Kaper, Manfred Heumann and Helge Ritter. *Improving transfer rates in brain computer interfacing: A case study.* Advances Neural Infor. Process. Syst. 15, pages 1107–1114, 2002.

[Mell 2008] Dominik Mell, Michael Bach and Sven P. Heinrich. *Fast stimulus sequences improve the efficiency of event-related potential P300 recordings.* J. Neurosc. Methods, vol. 174, no. 2, pages 259 – 264, 2008.

[Mellinger 2007] Jurgen Mellinger, Gerwin Schalk, Christoph Braun, Hubert Preissl, Wolfgang Rosenstiel, Niels Birbaumer and Andrea Kübler. *An MEG-based brain-computer interface (BCI).* NeuroImage, vol. 36, no. 3, pages 581 – 593, 2007.

[Mercure 2008] Evelyne Mercure, Frederic Dick and Mark H. Johnson. *Featural and configural face processing differentially modulate ERP components*. Brain Research, vol. 1239, pages 162 – 170, 2008.

[Millan 2004] Jd.R. Millan, F. Renkens, J. Mourino and W. Gerstner. *Noninvasive brain-actuated control of a mobile robot by human EEG*. IEEE Trans. Biomed. Eng., vol. 51, no. 6, pages 1026 –1033, june 2004.

[Monastra 2005] Vincent J. Monastra, Steven Lynn, Michael Linden, Joel F. Lubar, John Gruzelier and Theodore J. LaVaque. *Electroencephalographic Biofeedback in the Treatment of Attention-Deficit/Hyperactivity Disorder*. Applied Psychophys. and Biofeedback, vol. 30, pages 95–114, 2005.

[Müller-Gerking 1999] J. Müller-Gerking, G. Pfurtscheller and H. Flyvbjerg. *Designing optimal spatial filters for single-trial EEG classification in a movement task*. Clinical Neurophysiology, vol. 110, no. 5, pages 787–798, 1999.

[Müller 2004] Klaus-Robert Müller, Ricardo Vigario, Frank Meinecke and Andreas Ziehe. *Blind Source Separation Techniques for Decomposing Event Related Brain Signals*. Int. J. Bifurcation Chaos, vol. 14, no. 2, pages 773–791, 2004.

[Naeem 2006] M Naeem, C Brunner, R Leeb, B Graimann and G Pfurtscheller. *Seperability of four-class motor imagery data using independent components analysis*. J. Neural Eng., vol. 3, no. 3, pages 208–216, 2006.

[Nam 2009] C. S. Nam, Y. Kim Y. Jeon Y. Li and H. Yoon. *Usability of the P300 Speller: Towards a More Sustainable Brain-Computer Interface*. eMinds: Int. J. Human-Comput. Interaction, vol. I, no. 5, pages 111–125, 2009.

[Neuper 2006] Christa Neuper, Gernot R. Müller-Putz, Reinhold Scherer and Gert Pfurtscheller. *Motor imagery and EEG-based control of spelling devices and neuroprostheses*. In Christa Neuper and Wolfgang Klimesch, editors, Event-Related Dynamics of Brain Oscillations, volume 159 of *Progress in Brain Research*, pages 393 – 409. Elsevier, 2006.

[Neurosky 2011] Neurosky. *MindWave neuroeadset*. `http://www.neurosky.com/`, 2011. Accessed in 2011.

[NIH 2011] NIH. *Spinal Cord Injury: Hope Through Research*. `http://www.ninds.nih.gov/disorders/sci/`, 2011.

[Nijboer 2008] Femke Nijboer, Adrian Furdea, Ingo Gunst, Jürgen Mellinger, Dennis J. McFarland, Niels Birbaumer and Andrea Kübler. *An auditory brain-computer interface (BCI)*. J. Neurosc. Methods, vol. 167, no. 1, pages 43 – 50, 2008.

[Nijboer 2010] Femke Nijboer, Niels Birbaumer and Andrea Kubler. *The influence of psychological state and motivation on brain-computer interface performance in patients with amyotrophic lateral sclerosis - a longitudinal study.* Frontiers in Neurosc., vol. 4, no. 0, 2010.

[Oppenheim 1999] Alan V. Oppenheim, Ronald W. Schafer and John R. Buck. Discrete-time signal processing (2nd ed.). Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1999.

[Palmowski 1995] A. Palmowski, W. H. Jost, J. Prudlo J. Osterhage, B. Kasmann, K. Schimrigk and et al. *Eye movement in amyotrophic lateral sclerosis: a longitudinal study.* German J. Ophthal., vol. 4, pages 355–362, 1995.

[Patel 2005] SH Patel and PN Azzam. *Characterization of N200 and P300: Selected Studies of the Event-Related Potential.* Int. J. Medical Sciences, no. 2, pages 147–154, 2005.

[Peng 2005] Hanchuan Peng, Fuhui Long and C. Ding. *Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy.* IEEE Trans. Pattern Analysis Machine Intell., vol. 27, no. 8, pages 1226–1238, Aug. 2005.

[Pfurtscheller 1998] G. Pfurtscheller, N. Christa, A. Scholgl and K. Lugger. *Separability of EEG Signals Recorded During Right and Left Motor Imagery Using Adaptive Autoregressive Parameters.* IEEE Trans. Rehab. Eng., vol. 6, no. 3, pages 316–324, September 1998.

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

[Pfurtscheller 2005] G Pfurtscheller, G. Müller-Putz, J. Pfurtscheller and R. Rupp. *EEG-Based Asynchronous BCI Controls Functional Electrical Stimulation in a Tetraplegic Patient.* EURASIP J. Applied Signal Process., vol. 2005, no. 19, pages 3152–3155, 2005.

[Piccione 2006] F. Piccione, F. Giorgi, P. Tonin, K. Priftis, S. Giove, S. Silvoni and et al. *P300-based brain computer interface: Reliability and performance in healthy and paralysed participants.* Clinical Neurophysiology, vol. 117, no. 3, pages 531–537, March 2006.

[Pineda 2003] J.A. Pineda, D.S. Silverman, A. Vankov and J. Hestenes. *Learning to control brain rhythms: making a brain-computer interface possible.* IEEE Trans. Neural Syst. Rehab. Eng., vol. 11, no. 2, pages 181 –184, 2003.

[Pires 2002] Gabriel Pires and Urbano Nunes. *A wheelchair steered through voice commands and assisted by a reactive fuzzy-logic controller.* J. Intell. Robotic Syst., vol. 34, pages 301–314, July 2002.

[Pires 2007] G. Pires, U. Nunes and M. Castelo-Branco. *Single-Trial EEG classification of Movement Related Potential*. 10th Int. IEEE Conf. Rehab. Robotics, ICORR, pages 569–574, 2007.

[Pires 2008a] G. Pires, U. Nunes and M. Castelo-Branco. *Brain Computer Interface Approaches to Control Mobile Robotic Devices*. 11th Int. Conf. Climbing Walking Robots and the Support Technologies for Mobile Machines, CLAWAR, pages 489–496, 2008.

[Pires 2008b] Gabriel Pires, Miguel Castelo-Branco and Urbano Nunes. *Visual P300-based BCI to steer a wheelchair: A Bayesian approach*. 30th Annual Int. IEEE Conf., EMBC2008, pages 658–661, Aug. 2008.

[Pires 2009a] G. Pires, U. Nunes and M. Castelo-Branco. *P300 Spatial Filtering and Coherence-Based Channel Selection*. 4th Int. IEEE Conf. Neural Eng, NER09, pages 311–314, Apr 2009.

[Pires 2009b] Gabriel Pires. *Comparison of feature selectors for P300 classifcation*. Technical report, Institute for Systems and Robotics, 2009.

[Pires 2011a] G. Pires, M. Torres, N. Casaleiro, U. Nunes and M. Castelo-Branco. *Playing Tetris with non-invasive BCI*. 1st Int. IEEE Conf. on Serious Games and Applications for Health, SeGAH, 2011.

[Pires 2011b] Gabriel Pires, Urbano Nunes and Miguel Castelo-Branco. *Comparison of a Row-column Speller vs a Novel Lateral Single-character Speller: assessment of BCI for severe motor disabled patients*. Clinical Neurophysiology (In press), 2011. doi:10.1016/j.clinph.2011.10.040.

[Pires 2011c] Gabriel Pires, Urbano Nunes and Miguel Castelo-Branco. *GIBS block speller: toward a gaze independent P300-based BCI*. 33rd Annual Int. IEEE Conf., EMBC2011, pages 6360–6364, 2011.

[Pires 2011d] Gabriel Pires, Urbano Nunes and Miguel Castelo-Branco. *Statistical spatial filtering for a P300-based BCI: Tests in able-bodied, and patients with cerebral palsy and amyotrophic lateral sclerosis*. J. Neurosci. Methods, vol. 195, no. 2, pages 270–281, 2011.

[Platt 1999] John C. Platt. *Using Analytic QP and Sparseness to Speed Training of Support Vector Machines*. Advances in Neural Information Processing systems (NIPS), vol. 11, pages 557–563, 1999.

[Polich 1996] John Polich, Patricia Crane Ellerson and Jill Cohen. *P300, stimulus intensity, modality, and probability*. Int. J. Psychophysiology, vol. 23, no. 1-2, pages 55 – 62, 1996.

[Polich 2007] John Polich. *Updating P300: An integrative theory of P3a and P3b*. Clinical Neurophysiology, vol. 118, no. 10, pages 2128 – 2148, 2007.

[Pope 2004] Alan Pope and Olafur Palsson. *Helping Video Games 'Rewire Our Minds'.* NASA technical report, 2004.

[Rakotomamonjy 2008] A. Rakotomamonjy and V. Guigue. *BCI Competition III: Dataset II- Ensemble of SVMs for BCI P300 Speller.* IEEE Trans. Biomed. Eng., vol. 55, no. 3, pages 1147–1154, March 2008.

[Ramoser 2000] H. Ramoser, J. Müller-Gerking and G. Pfurtscheller. *Optimal spatial filtering of single trial EEG during imagined hand movement.* IEEE Trans. Rehab. Eng., vol. 8, no. 4, pages 441–446, Dec 2000.

[Rebsamen 2010] B. Rebsamen, Guan Cuntai, Zhang Haihong, Wang Chuanchu, Teo Cheeleong, M.H. Ang and E. Burdet. *A Brain Controlled Wheelchair to Navigate in Familiar Environments.* IEEE Trans. Neural Syst. Rehab. Eng., vol. 18, no. 6, pages 590–598, Dec 2010.

[Reuderink 2008] B. Reuderink. *Games and Brain-Computer Interfaces: The State of the Art.* Technical report TR-CTIT-08-81, Centre for Telematics and Infor. Tech., University of Twente, Enschede, 2008.

[Ribeiro 2009] A. Ribeiro, A. Sirgado, J. Aperta, A. Lopes, J. Guilherme, P. Correia, G. Pires and U. Nunes. *A Low-Cost EEG Stand-Alone Device for Brain Computer Interface.* 2nd Int. Joint Conf. Biomed. Eng. Syst. Technologies, Biodevices, pages 430–433, 2009.

[Rish 2001] I. Rish. *An empirical study of the naive Bayes classifier.* IJCAI-01 workshop, 2001.

[Rivet 2009] B. Rivet, A. Souloumiac, V. Attina and G. Gibert. *xDAWN Algorithm to Enhance Evoked Potentials: Application to Brain-Computer Interface.* IEEE Trans. Biomed. Eng., vol. 56, no. 8, pages 2035 –2043, aug. 2009.

[Rohrbaugh 1974] John Rohrbaugh, Emanuel Donchin and Charles Eriksen. *Decision making and the P300 component of the cortical evoked response.* Attention, Perception, amp and Psychophysics, vol. 15, pages 368–374, 1974.

[Salvaris 2009] M Salvaris and F Sepulveda. *Visual modifications on the P300 speller BCI paradigm.* J. Neural Eng., vol. 6, no. 4, 2009.

[Sampaio 2011] Joana Sampaio, Elzbieta Bobrowicz-Campos, Rui Andre, Inês Almeida, Pedro Faria, Cristina Januario, Antonio Freire and Miguel Castelo-Branco. *Specific impairment of visual spatial covert attention mechanisms in Parkinson's disease.* Neuropsychologia, vol. 49, no. 1, pages 34 – 42, 2011.

[Sanei 2007] S. Sanei and J. A. Chambers. EEG Signal Processing. John Wiley & Sons, 2007.

[Schalk 2004] G. Schalk, D.J. McFarland, T. Hinterberger, N. Birbaumer and J.R. Wolpaw. *BCI2000: a general-purpose brain-computer interface (BCI) system.* IEEE Trans. Biomed. Eng., vol. 51, no. 6, pages 1034–1043, June 2004.

[Schalk 2008] G Schalk, K J Miller, N R Anderson, J A Wilson, M D Smyth, J G Ojemann, D W Moran, J R Wolpaw and E C Leuthardt. *Two-dimensional movement control using electrocorticographic signals in humans.* J. Neural Eng., vol. 5, no. 1, page 75, 2008.

[Scherer 2004] R. Scherer, G.R. Müller, C. Neuper, B. Graimann and G. Pfurtscheller. *An asynchronously controlled EEG-based virtual keyboard: improvement of the spelling rate.* IEEE Trans. Biomed. Eng., vol. 51, no. 6, pages 979 –984, june 2004.

[Schlögl 2007] Alois Schlögl, Julien Kronegg, Jane E. Huggins and Steve G. Mason. Evaluation Criteria for BCI Research, chapter 19, pages 327–342. MIT Press, 2007.

[Scolari 2007] Miranda Scolari, Andrew Kohnen, Brian Barton and Edward Awh. *Spatial attention, preview, and popout: Which factors influence critical spacing in crowded displays?* J. Vision, vol. 7, no. 2, 2007.

[Sekihara 2001] K. Sekihara, S.S. Nagarajan, D. Poeppel, A. Marantz and Y. Miyashita. *Reconstructing spatio-temporal activities of neural sources using an MEG vector beamformer technique.* IEEE Trans. Biomed. Eng., vol. 48, no. 7, pages 760 –771, july 2001.

[Sellers 2006a] Eric W. Sellers, Dean J. Krusienski, Dennis J. McFarland, Theresa M. Vaughan and Jonathan R. Wolpaw. *A P300 event-related potential brain-computer interface (BCI): The effects of matrix size and inter stimulus interval on performance.* Biological Psychology, vol. 73, no. 3, pages 242 – 252, 2006.

[Sellers 2006b] Eric W. Sellers, Dean J. Krusienski, Dennis J. McFarland, Theresa M. Vaughan and Jonathan R. Wolpaw. *A P300 event-related potential brain-computer interface (BCI): The effects of matrix size and inter stimulus interval on performance.* Biological Psychology, vol. 73, no. 3, pages 242 – 252, 2006.

[Serby 2005] H. Serby, E. Yom-Tov and G. Inbar. *An Improved P300-Based Brain-Computer Interface.* IEEE Trans. Neural Syst. Rehab. Eng., vol. 13, pages 89–98, March 2005.

[Shannon 1948] Claude E. Shannon. *A mathematical theory of Communication.* The Bell system technical journal, vol. 27, pages 379–423, July 1948.

[Silva 2008] Maria Fatima Silva, Susana Maia-Lopes, Catarina Mateus, Manuela Guerreiro, Joana Sampaio, Pedro Faria and Miguel Castelo-Branco. *Retinal and cortical patterns of spatial anisotropy in contrast sensitivity tasks.* Vision Research, vol. 48, pages 127–135, 2008.

[Silva 2010] MF Silva, C Mateus, A Reis, S Nunes, P Fonseca and M. Castelo-Branco. *Asymmetry of visual sensory mechanisms: electrophysiological, structural, and psychophysical evidences.* J. Vision, vol. 10, 2010.

[Soong 1995] A.C.K. Soong and Z.J. Koles. *Principal-component localization of the sources of the background EEG.* IEEE Trans. Biomed. Eng., vol. 42, no. 1, pages 59–67, Jan. 1995.

[Srinivasan 1999] R. Srinivasan. *Methods to improve the spatial resolution of EEG.* Int. J. Bioelectromagnetism, vol. 1, pages 102–111, 1999.

[Sutter 1992] Erich E. Sutter. *The brain response interface: communication through visually-induced electrical brain responses.* J. Microcomputer Applications, vol. 15, no. 1, pages 31 – 45, 1992.

[Takano 2009] Kouji Takano, Tomoaki Komatsu, Naoki Hata, Yasoichi Nakajima and Kenji Kansaku. *Visual stimuli for the P300 brain-computer interface: A comparison of white/gray and green/blue flicker matrices.* Clinical Neurophysiology, vol. 120, no. 8, pages 1562 – 1566, 2009.

[Tanaka 2005] K. Tanaka, K. Matsunaga and H.O. Wang. *Electroencephalogram-Based Control of an Electric Wheelchair.* IEEE Trans. Robotics, vol. 21, no. 4, pages 762–766, Aug 2005.

[Tangermann 2008] Michael Tangermann, Matthias Krauledat, Konrad Grzeska, Max Sagebaum, Benjamin Blankertz and Klaus-Robert Müller. *Playing pinball with non-invasive BCI.* Advances Neural Inform. Process. Syst., vol. 21, 2008.

[Taylor 2002] Dawn M. Taylor, Stephen I. Helms Tillery and Andrew B. Schwartz. *Direct Cortical Control of 3D Neuroprosthetic Devices.* Science, vol. 296, no. 5574, pages 1829–1832, 2002.

[Teixeira 2010] M Teixeira, M Castelo-Branco, S Nascimento and V Almeida. *The P300 signal is monotonically modulated by target saliency level irrespective of the visual feature domain.* Acta Ophthalmologica, vol. 88, pages 0–0, 2010.

[Teixeira 2011] M. Teixeira, G. Pires, M. Raimundo, S.M. Nascimento, V.M. de Almeida and M. Castelo-Branco. *Neural bases of visual attention and saliency: detecting feature independent perceptual saliency at a single trial level using oddball EEG signals.* IBILI Meeting 2011, 2011.

[Thulasidas 2006] M. Thulasidas, Cuntai Guan and Jiankang Wu. *Robust classification of EEG signal for brain-computer interface.* IEEE Trans. Neural Syst. Rehab. Eng., vol. 14, no. 1, pages 24–29, March 2006.

[Tomioka 2006] Ryota Tomioka, Guido Dornhege, Guido Nolte, Benjamin Blankertz, Kazuyuki Aihara and Klaus-Robert Müller. *Spectrally Weighted Common Spatial Pattern Algorithm for Single Trial EEG Classification.* Technical report 40, Dept. of Mathematical Engineering, The University of Tokyo, July 2006.

[Tomioka 2007] Ryota Tomioka, Kazuyuki Aihara and Klaus-Robert Müller. *Logistic Regression for Single Trial EEG Classification.* In B. Schölkopf, J. Platt and T. Hoffman, editors, Advances in Neural Information Processing Systems 19, pages 1377–1384. MIT Press, 2007.

[Torres 2011] Mario Torres. Controlo do jogo tetris através de uma interface cérebro computador usando sinais electroencefalográficos. Msc dissertation, Universidade de Coimbra - Departamento de Engenharia Electrotecnica, 2011.

[Townsend 2004] G. Townsend, B. Graimann and G. Pfurtscheller. *Continuous EEG classification during motor imagery-simulation of an asynchronous BCI.* IEEE Trans. Neural Syst. Rehab. Eng., vol. 12, no. 2, pages 258 –265, june 2004.

[Townsend 2010] G. Townsend, B.K. LaPallo, C.B. Boulay, D.J. Krusienski, G.E. Frye, C.K. Hauser and et al. *A novel P300-based brain-computer interface stimulus presentation paradigm: Moving beyond rows and columns.* Clinical Neurophysiology, vol. 121, no. 7, pages 1109 – 1120, 2010.

[Treder 2010] Matthias Treder and Benjamin Blankertz. *(C)overt attention and visual speller design in an ERP-based brain-computer interface.* Behav. Brain Funct., vol. 6, no. 1, 2010.

[Trees 2002] H. Van Trees. Optimum array processing. part IV of detection, estimation and modulation theory. Wiley Interscience, 2002.

[Van Veen 1988] Barry D. Van Veen and Kevin M. Buckley. *Beamforming: A Versatile Approach to Spatial Filtering.* IEEE Signal Processing Magazine, vol. 5, no. 2, pages 4–24, April 1988.

[Van Veen 1997] B.D. Van Veen, W. Van Drongelen, M. Yuchtman and A. Suzuki. *Localization of brain electrical activity via linearly constrained minimum variance spatial filtering.* IEEE Trans. Biomed. Eng., vol. 44, no. 9, pages 867 –880, sept. 1997.

[Vapnik 1995] Vladimir N. Vapnik. The nature of statistical learning theory. Springer-Verlag New York, Inc., New York, NY, USA, 1995.

[Vidal 1973] J Vidal. *Toward Direct Brain-Computer Communication.* Annual Review Biophysics Bioeng., vol. 2, pages 157–180, 1973.

[Vidaurre 2011] C. Vidaurre, M. Kawanabe, P. von Bunau, B. Blankertz and K.R. Müller. *Toward Unsupervised Adaptation of LDA for Brain-Computer Interfaces.* IEEE Trans. Biomed. Eng., vol. 58, no. 3, pages 587–597, march 2011.

[Vuckovic 2008] A. Vuckovic and F. Sepulveda. *Delta band contribution in cue based single trial classification of real and imaginary wrist movements.* Medical Biological Eng. Computing, vol. 46, no. 6, pages 529–539, June 2008.

[Wang 2006] Yijun Wang, Ruiping Wang, Xiaorong Gao, Bo Hong and Shangkai Gao. *A practical VEP-based brain-computer interface.* IEEE Trans. Neural Syst. Rehab. Eng., vol. 14, no. 2, pages 234 –240, june 2006.

[Wang 2011] Qiang Wang, Olga Sourina and Minh Nguyen. *Fractal dimension based neurofeedback in serious games.* The Visual Computer, vol. 27, pages 299–309, 2011.

[Weiskopf 2004] N. Weiskopf, K. Mathiak, S.W. Bock, F. Scharnowski, R. Veit, W. Grodd, R. Goebel and N. Birbaumer. *Principles of a brain-computer interface (BCI) based on real-time functional magnetic resonance imaging (fMRI).* IEEE Trans. Biomed. Eng., vol. 51, no. 6, pages 966 –970, june 2004.

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

[Wolpaw 2002] Jonathan R Wolpaw, Niels Birbaumer, Dennis J McFarland, Gert Pfurtscheller and Theresa M Vaughan. *Brain-computer-interface for communication and control.* Clinical Neurophysiology, vol. 113, no. 6, pages 767–791, 2002.

[Wolpaw 2004] Jonathan R. Wolpaw and Dennis J. McFarland. *Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans.* Proc. National Academy Sciences of the United States of America, vol. 101, no. 51, pages 17849–17854, 2004.

[Wu 2008] Wei Wu, Xiaorong Gao, Bo Hong and Shangkai Gao. *Classifying Single-Trial EEG During Motor Imagery by Iterative Spatio-Spectral Patterns Learning (ISSPL).* IEEE Trans. Biomed. Eng., vol. 55, no. 6, pages 1733 –1743, june 2008.

[Xu 2004] Neng Xu, Xiaorong Gao, Bo Hong, Xiaobo Miao, Shangkai Gao and Fusheng Yang. *BCI competition 2003-data set IIb: enhancing P300 wave detection using ICA-based subspace projections for BCI applications.* IEEE Trans. Biomed. Eng., vol. 51, no. 6, pages 1067 –1072, june 2004.

[Zhang 2008] Haihong Zhang, Cuntai Guan and Chuanchu Wang. *Asynchronous P300-Based Brain–Computer Interfaces: A Computational Approach With Statistical Models.* IEEE Trans. Biomed. Eng., vol. 55, no. 6, pages 1754–1763, June 2008.