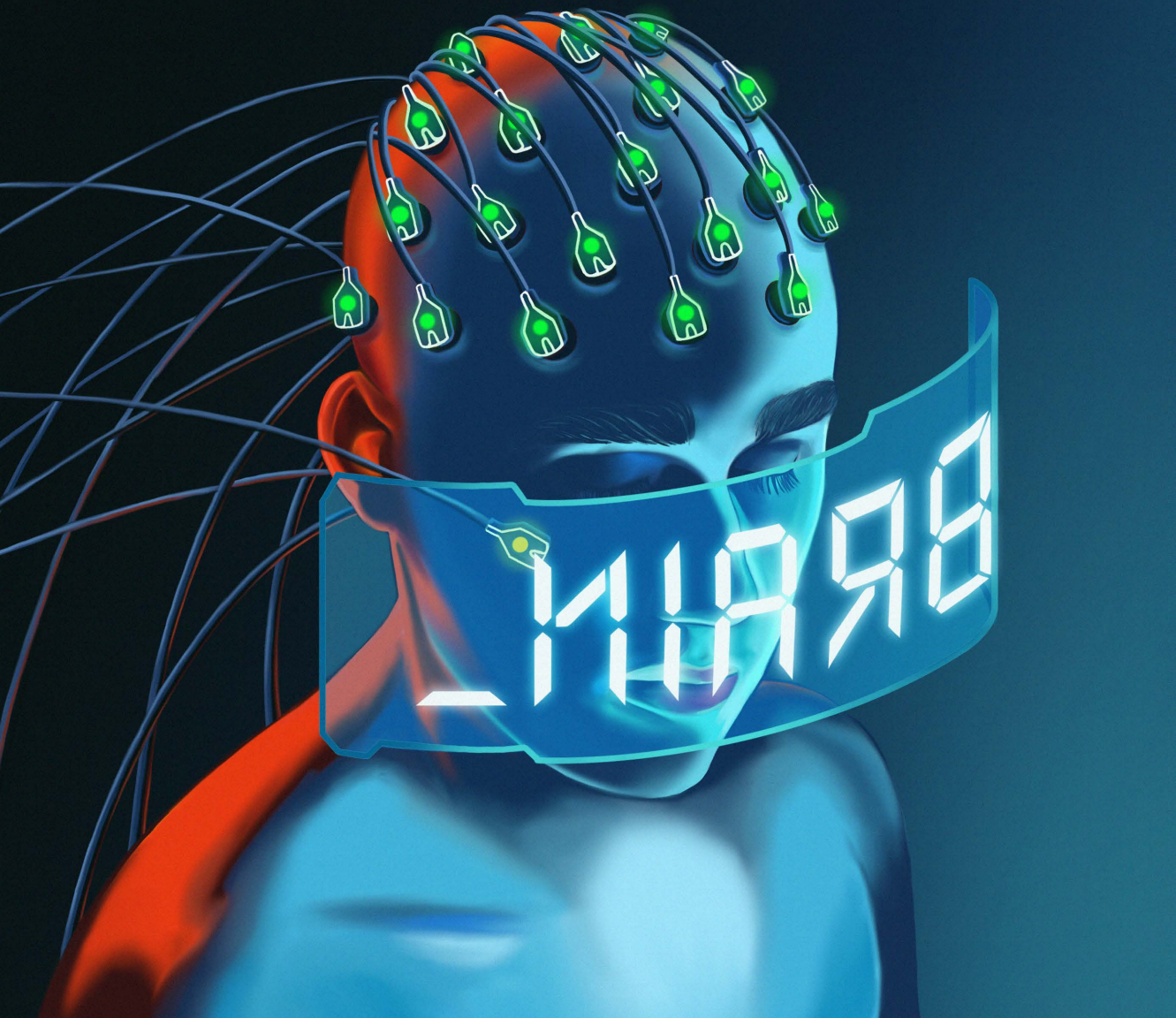
Toward Practical P300-based Brain–Computer Interfaces

Asynchrony, Channel Selection and Assistive Applications

Víctor Martínez-Cagigal

Doctoral Thesis



Universidad de Valladolid

Doctoral Program of
Information and Telecommunication Technologies

Doctoral Thesis

Toward Practical P300-based Brain–Computer Interfaces: Asynchrony, Channel Optimization and Assistive Applications

Thesis presented by **D. Víctor Martínez-Cagigal**
to apply for the *Ph.D. degree*
from the *University of Valladolid*

Directed by:
Dr. Roberto Hornero Sánchez

2020
Valladolid, Spain

A mi familia,

Defense

Título	Toward Practical P300-based Brain–Computer Interfaces: Asynchrony, Channel Optimization and Assistive Applications
Autor	D. Víctor Martínez-Cagigal
Director	Dr. D. Roberto Hornero Sánchez
Departamento	Teoría de la Señal y Comunicaciones e Ingeniería Telemática

---

Tribunal

Presidente	Dr. D. José María Azorín Poveda
Vocal	Dr. D. Jorge Leon Morales Quezada
Secretario	Dr. D. Jesús Poza Crespo

acuerda otorgarle la calificación de

En Valladolid, a de del

Universidad de Valladolid

Escuela Técnica Superior de Ingenieros de Telecomunicación
Dpto. de Teoría de la Señal y Comunicaciones e Ingeniería Telemática

Research Stay for the International Mention

City:	Graz (Austria)
Faculty:	Graz University of Technology
Institution:	Institute of Neural Engineering
Research group:	Laboratory of Brain–Computer Interfaces
Dates:	09/09/2019-10/12/2019
Duration:	92 days (3 months)
Supervisor:	Prof. Dr. Gernot Müller-Putz

GRAZ BCI

Agradecimientos

En primer lugar, quisiera expresar mi más sincera gratitud al Dr. Roberto Hornero Sánchez por guiarme durante todos estos años en mi odisea investigadora. Parece mentira que hayan pasado seis años desde aquel día que solicité tu supervisión para realizar un trabajo fin de grado sobre sistemas *brain–computer interface* (BCI), y seis años también desde que me advertiste por vez primera que realizar un doctorado es un proceso arduo y sacrificado. Tus consejos, tu paciencia y tu compromiso han sido esenciales para poder concluir este periodo felizmente.

Me gustaría hacer extensivo este agradecimiento al resto del Grupo de Ingeniería Biomédica (GIB), a los que están y a los que estuvieron, a Carlos, Jesús, María, Daniel, Gonzalo, Rebeca, Javier G., Alejandro, Luis F., Fernando, Pablo, Verónica, Saúl, Roberto, Eduardo, Adrián, Víctor R., Aarón, Jorge, Javier O., Víctor G. y Marcos. El GIB no solo es el seno de una investigación de prestigio internacional, sino también un lugar con un ambiente distendido donde da gusto trabajar. Las ideas locas derivadas de los '*happy meals*', los partidos, las quedadas y los viajes han hecho de este ciclo una etapa inolvidable. Con especial cariño quiero agradecer a Javier G. por sus constantes consejos y la paciencia que ha demostrado resolviendo mis dudas desde el principio hasta el final. Te has convertido en lo más parecido a un *senpai*, como dicen en el país de sol naciente. Tampoco puedo olvidarme de Eduardo, compañero de batallas por excelencia. Sin tu compañía todos los registros, reuniones, ruedas de prensa y problemas técnicos habrían sido mucho más difíciles, gracias.

I would also like to thank Dr. Gernot Müller-Putz and the Institute of Neural Engineering for embracing me in one of the best BCI research centers from all over the world. Undoubtedly, my research internship in Graz was an unforgettable experience, where I grew up professionally and personally. I would especially like to thank Reinmar and Valeria for their invaluable help with my study; as well as to Andreea, Philipp, Luka, Joana, Nitikorn, Aleksandra, Catarina, and Lea to

make me feel like home.

El mayor de los agradecimientos se lo dedico a mi familia, mis padres y mi hermana, por apoyarme, educarme, encarrilarme, por darme la oportunidad de poder estudiar lo que me apasiona y por inculcarme los principios de los que me enorgullezco. Sin duda, sois la piedra angular que ha permitido que alcance este hito, os lo agradezco.

Finalmente, me gustaría concluir estas líneas con mi más profundo agradecimiento a Giselle, la persona más luchadora y valiente que conozco. Me has enseñado el poder de las emociones, la empatía, la subjetividad, el altruismo y el respeto. Te doy las gracias por complementarme, por aguantarme y por haberme apoyado en todo momento: en la cercanía y en la distancia, en la salud y en la enfermedad. Sin tu ayuda este camino hubiese sido más difícil, muchas gracias.

Abstract

For decades, mankind has fantasized about the possibility of controlling devices with our minds. Despite there is still a long way to go to achieve that goal, recent progresses in neuroscience took a step forward and contributed to the development of the first brain–computer interface (BCI) systems. Through the analysis of electroencephalographic (EEG) signals, BCIs are able to decode users' intentions into application commands. Due to its ability to enhance or even replace nervous system outputs, BCIs have emerged as novel assistive technologies that could improve the quality of life of the severely disabled. Nevertheless, these systems currently do not provide the required reliability to take the leap from laboratories to real environments. Among the problems that current BCIs should face to, poor performances, need of supervision and lack of portability and validation with target users stand out.

In this context, the present Doctoral Thesis is focused on the development of novel signal processing methodologies and assistive applications that contribute to provide a real use of BCIs by motor-disabled people. From all kinds of BCI control signals, the studies included in this compendium of publications use the P300 evoked potential due to its versatility and reliability. In fact, BCIs that are currently used by disabled people on a daily basis are generally P300-based, restricting many other control signals to purely academic domains.

The contributions of this study are canalized in three different ways. Firstly, two asynchronous algorithms are proposed. By default, P300-based BCIs are synchronous systems, presenting an inability to monitor users' attention and causing unintended command selections even when users are not paying attention to the *oddball* stimulation. This mode is unpractical and should be avoided whenever a BCI is intended to be feasible in a real setup. Here, we propose (i) a thresholding wrapper approach, which discriminates between control (i.e., attending) and non-control (i.e., ignoring) states based on the classifier scores; and (ii) a fil-

ter entropy-based approach, which allows to characterize EEG signals from both states and extract discriminative information independently of the classifier. The thresholding approach was integrated in two different assistive applications and tested with motor-disabled users, while the entropy-based approach was studied offline with healthy subjects.

Secondly, we propose the application of evolutionary single and multi-objective meta-heuristics to select optimal channel sets for each user. This optimization is often overlooked in the BCI literature due to its inherent complexity. However, it is beneficial to improve the performance and users' comfort, as well as to reduce power consumption and the cost of the system. Furthermore, we present a novel multi-objective algorithm especially designed for the BCI framework, the dual-front genetic algorithm (DFGA). They were tested with three public databases that recorded data of healthy subjects from different *oddball* paradigms.

Lastly, we present the design, development and evaluation of two novel asynchronous assistive BCI applications: (i) a web browser, and (ii) a social networking app for smartphones. The web browser is intended to be controlled using a laptop, selecting page links through a node tagging approach via row-col paradigm matrices. The social networking app, by contrast, allows users to control Twitter and Telegram in their smartphones. Both were evaluated with a population of motor-disabled users in order to assess their feasibility in a real setup, detailing not only quantitative measures (e.g., accuracy, timings), but also their qualitative opinions and suggestions using questionnaires.

Our findings showed that the integration of an asynchronous management significantly improved the performance in assistive BCIs. Particularly, it was found that control signals are more complex and irregular than non-control ones, allowing a reliable monitoring of users' attention using entropy-based metrics (up to 94.4% accuracy). These outcomes brought to light the need of implementing an asynchronous stage to provide a comprehensive control of the BCI and thus, to support the personal autonomy of the target users. Moreover, results showed that optimal channel sets present a high inter-subject variability, making the channel selection stage essential to optimize the overall performance of each user. In that sense, the balanced combination of deterministic and stochastic strategies of DFGA fostered the overcoming of existing algorithms in terms of performance and convergence, allowing the supervisor to select an optimal set in function of the number of channels available. Concerning the assistive applications, results showed that the motor-disabled participants obtained significantly lower accuracies than healthy subjects (web browser: 95.8%, social networking: 92.3%), which

states that a validation with target users is crucial to assure the feasibility of BCIs in a real context. Despite this fact, the performances of motor-disabled users were more than enough to claim the viability of both applications (web browser: 84.1%, social networking: 80.6%), likely due to their asynchronous nature. Furthermore, participants stated that they could imagine themselves using both applications on their daily basis. We feel that these studies will contribute to move toward a real use of these systems by motor-disable people, aiming at improving their personal autonomy and quality of life.

Acronyms

A1	Primary auditory cortex
ACO	Ant Colony Optimization
ADC	Analog-to-Digital Converter
ALS	Amyotrophic Lateral Sclerosis
AUC	Area Under ROC Curve
BCI	Brain–Computer Interface
BA	Bees Algorithm
BAS	Binary Ant System
(B)DE	(Binary) Differential Evolution
BE	Backward Elimination
BH	Benjamini-Hochberg correction
BOLD	Blood-Oxygen-Level-Dependent response
BP	Bereitschaftspotential
(B)PSO	(Binary) Particle Swarm Optimization
CAR	Common Average Reference
CCA	Canonical Correlation Analysis
CNS	Central Nervous System
CNV	Contingent Negative Variation
CP	Cerebral Palsy
CS	Center Speller paradigm
DFGA	Dual-Front Sorting Genetic Algorithm
EEG	Electroencephalography
ERD	Event-Related Desynchronization
ERP	Event-Related Potential
ERS	Event-Related Synchronization
FA	Firefly Algorithm
FDR	False Discovery Rate

FIR	Finite Impulse Response
FLD	Fisher's Linear Discriminant
FIRDA	Frontal Intermittent Rhythmic Delta Activity
FRDA	Friedreich's Ataxia
fMRI	Functional Magnetic Resonance Imaging
fNIRS	Functional Near-Infrared Spectroscopy
FNR	False Negative Rate
FS	Forward Selection
FWER	Family-Wise Error Rate
GA	Genetic Algorithm
HS	Healthy Subjects
IIR	Infinite Impulse Response
ISI	Inter-Stimuli Interval
JCR	Journal Citation Reports
LIS	Locked-In Syndrome
LDA	Linear Discriminant Analysis
LOO	Leave one out
M1	Primary Motor Cortex
MD	Muscular Dystrophy
MDS	Motor-Disabled Subjects
MEG	Magnetoencephalography
MI	Motor Imagery
MOBDE	Multi-Objective Binary Differential Evolution
MOPSO	Multi-Objective Particle Swarm Optimization
MRCP	Movement-Related Cortical Potential
MS	Multiple Sclerosis
MSE	Multiscale Entropy
NF	Neurofeedback
NSGA-II	Non-Sorting Genetic Algorithm 2
OCM	Output Characters per Minute
OIRDA	Occipital Intermittent Rhythmic Delta Activity
PET	Positron Emission Tomography
PFC	Prefrontal Cortex
PLS	Partial Least Squares
PSD	Power Spectral Density
RCP	Row-Col Paradigm
ROC	Receiving Operating Characteristic

RSVP	Rapid Serial Visual Presentation
S1	Primary Somatosensory Cortex
SampEn	Sample Entropy
SCI	Spinal Cord Injury
SCP	Slow Cortical Potentials
SD	Stimulus Duration
SMR	Sensorimotor Rhythms
SOA	Stimulus Onset Asynchrony
SPEA2	Strength Pareto Evolutionary Algorithm 2
SQUID	Superconducting Quantum Interference Device
SQ-UKF	Square-Root Unscented Kalman Filter
SSVEP	Steady-State Visual Evoked Potential
SVM	Support Vector Machine
SW	Step-Wise regression
V1	Primary Visual Cortex
VEP	Visual Evoked Potential

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The current Doctoral Thesis focuses on contributing the Brain–Computer Interface (BCI) field by designing, developing and testing novel signal processing methods and assistive applications. BCI systems, which allow users to control applications or external devices by processing their own brain signals, could rely on either invasive or non-invasive approaches. Owing to the practicality of the system, this study has been focused on non-invasive BCIs that analyze electroencephalographic (EEG) signals. This research have led to the publication of a total of four articles in journals indexed in the Journal Citation Reports (JCR) from the Web of Science™. In particular, up to three papers were published between September 2017 and April 2019. Additionally, a fourth article is currently under review. As a result of the aforementioned scientific productivity, this work is written as a compendium of publications.

The thematic consistency of the manuscripts in relation to the Doctoral Thesis is proved in section 1.1. The general context is briefly described in section 1.2, which introduces biomedical engineering, signal processing and evolutionary computation fields. Section 1.3 is devoted to the human brain, where the cerebrum anatomy, its functions, and the measurement of brain activity are discussed. Section 1.4 is focused on providing the basis for understanding BCI systems, including their structure, invasiveness, and the control signals that correlate with the users' intentions. Current limitations of the BCI systems are detailed as well. The latter, indeed, are the cause motivating the research problem. Due to the ability of BCIs to establish a communication pathway between our brain and the environment, these systems are specially suited to improve the quality of life of the severely disabled. Therefore, motor disabilities of the target users that evaluated

the applications included in the current doctoral thesis are described in section 1.5. Finally, section 1.6 provides a comprehensive state-of-the-art revision of previous approaches.

1.1 Compendium of publications: thematic consistency

The idea of establishing a link between our brains and the environment, as well as the possibility of acting through neural signals rather than muscles, have fascinated mankind for a century. The discovering of EEG by Hans Berger in 1929 (Berger, 1929), the initial work of Jacques Vidal in 1973 and further progresses in neuroscience made this idea even more plausible (Vidal, 1973a). Nowadays, science fiction is becoming reality. For the past 25 years, research groups have made great efforts to decode neural signals and provide different BCI systems to disabled users and general population. Invasive and non-invasive, dependent or independent, exogenous or endogenous, active or passive, synchronous or asynchronous; many types of BCIs have arisen and are being improved by the research community. Nonetheless, most BCI systems still lack the required reliability for emancipating from the laboratories to a practical environment.

Hardware limitations, scarce performances, lack of validation or need of supervision are some of the problems that current BCIs should face up to. In this context, the present Doctoral Thesis aims at developing novel signal processing methodologies and assistive applications, both intended to provide a real use of the BCIs by motor-disabled people. Particularly, all the studies that have been carried out in this work deal with P300 evoked potentials, a reliable control signal that provides such performances that P300-based BCIs are the only ones that are currently used by disabled people on a daily basis (Sellers Eric W., 2012). This is the common thread shared by all the papers included in the present compendium of publications. The relationship of the papers with each other and with the aforementioned objective is illustrated in the Figure 1.1, whose main contributions are depicted inside the common structure of a BCI system.

According to this non-chronological order, the first two papers were focused on the signal processing level, whereas the last two were aimed at providing assistive applications. The first paper (Martínez-Cagigal et al., 2019b) dealt with the need of supervision of current BCIs. Nowadays, most P300-based BCIs rely on synchronous paradigms, leading to a random selection of commands even when

users are not paying attention to the stimuli. We investigated if sample entropy features are able to track users' attention in real-time and provide an asynchronous (i.e., self-paced) control of the system. We also developed a wrapper threshold-based approach, which was then applied in the third (Martínez-Cagigal et al., 2017) and fourth (Martínez-Cagigal et al., 2019a) papers. In the second paper (Martínez-Cagigal et al., 2020), we assessed the ability of three single- and three multi-objective meta-heuristics to select a customized channel set for each user. Then, we developed and proposed a novel method to overcome their limitations, and we established a set of guidelines for adapting any meta-heuristic algorithm to the P300-BCI channel selection problem. The last two papers were intended to develop practical assistive BCIs for bridging the accessibility gap in new technologies for the severely disabled. In particular, the third paper (Martínez-Cagigal et al., 2017) was aimed at providing an asynchronous P300-based BCI web browser. For the sake of viability, the system was tested with sixteen multiple sclerosis patients and five healthy volunteers, reaching average accuracies of 84.14% and 95.75%, respectively. Lastly, the fourth paper (Martínez-Cagigal et al., 2019a) was also intended to provide a P300-based assistive application for motor-disabled people. This time the paper presented an asynchronous BCI for controlling smartphone-based social networks. The system was tested with eighteen motor-disabled and ten healthy subjects, achieving mean accuracies of 80.6% and 92.3%, respectively.

Due to the structure of the present Doctoral Thesis, organized as a compendium of publications, consulting each paper separately is essential for a comprehensive understanding of this document as a whole. Therefore, Appendix A includes the aforementioned manuscripts. Titles, authors, and abstracts of each one, as well as the journals in which they were published are shown below:

Asynchronous control of P300-based Brain–Computer Interfaces using sample entropy (Martínez-Cagigal et al., 2019b).

Víctor Martínez-Cagigal, Eduardo Santamaría-Vázquez and Roberto Hornero. *Entropy*, vol. 21(3), p. 230, 2019. Impact factor in 2019: 2.494, Q2 in "PHYSICS, MULTIDISCIPLINARY" (JCR-WOS).

Abstract: Brain-computer interfaces (BCI) have been traditionally worked using synchronous paradigms. In recent years, many efforts have been made to reach an asynchronous management, providing users with the ability to decide when a command should be selected. However, to the best of our knowledge, entropy metrics have not yet been explored. The present study follows a twofold purpose:

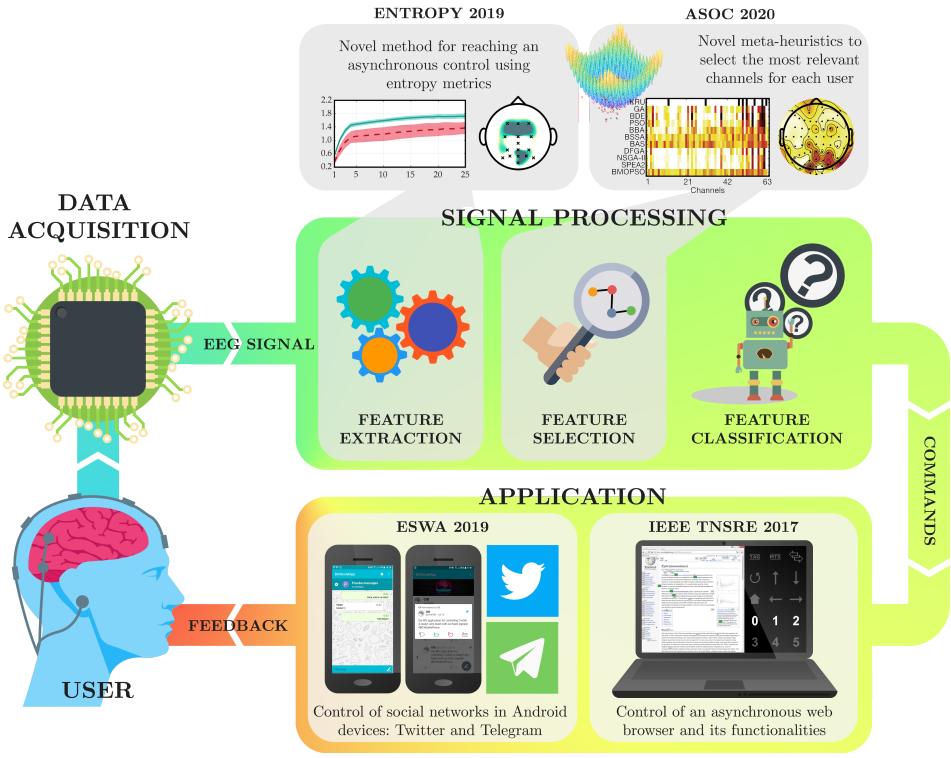


Figure 1.1: Main contributions of the papers included in the compendium of publications, arranged along the common structure of a BCI system. ASOC: Applied Soft Computing, IEEE TNSRE: IEEE Transactions on Neural Systems and Rehabilitation Engineering, ESWA: Expert Systems With Applications.

(i) to characterize both control and non-control states by examining the regularity of the EEG signals; and (ii) to assess the efficacy of a scaled version of the sample entropy algorithm to provide an asynchronous control for BCI systems. Ten healthy subjects participated in the study, who were asked to spell words through a visual oddball-based paradigm, attending (i.e., control) and ignoring (i.e., non-control) the stimuli. An optimization stage was performed for determining a common combination of hyperparameters for all subjects. Afterwards, these values were used to discern between both states using a linear classifier. Results show that control signals are more complex and irregular than non-control ones, reaching an average accuracy of 94.40% in classification. In conclusion, the present study demonstrates that the proposed framework is useful to monitor the attention of the user and grant the asynchrony of the BCI system.

Brain–Computer Interface channel selection optimization using meta-heuristics and evolutionary algorithms (Martínez-Cagigal et al., 2020).

Víctor Martínez-Cagigal, Eduardo Santamaría-Vázquez and Roberto Hornero. *Applied Soft Computing*, Under Review (R2), 2020. Impact factor in 2019 (last year available): 5.472, D1 (Q1) in "COMPUTER SCIENCE, INTERDISCIPLINARY APPLICATIONS", and Q1 in "COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE" (JCR-WOS).

Many brain–computer interface (BCI) studies overlook the channel optimization procedure due to its inherent complexity. However, a suitable channel selection procedure increases the performance and users' comfort while reducing the cost of the system. Evolutionary meta-heuristics, which have gained importance due to their excellent performances in solving complex problems, have not been fully exploited yet in this context. The purpose of the study is two-fold: (1) to propose a novel algorithm to find an optimal channel set for each user and compare its usefulness with other existing meta-heuristics; and (2) to establish guidelines for adapting these optimization strategies to this framework. A total of 3 single-objective (GA, BDE, BPSO) and 3 multi-objective (NSGA-II, BMOPSO, SPEA2) existing algorithms have been adapted and tested with 3 public databases: 'BCI competition III–dataset II', 'Center Speller' and 'RSVP Speller'. Dual-Front Sorting Algorithm (DFGA), a novel multi-objective discrete method especially adapted to the BCI framework, is proposed as well. Results show that all the applied meta-heuristics reached accuracies that significantly outperformed the entire set and the common 8-channel set of P300-based BCIs. DFGA showed a significant improvement of accuracy of 3.9% over the 8-channel set using also 8 channels; and obtained similar accuracies than it using a mean of only 4.66 channels. Binary-based algorithms stood out because of their faster convergence, especially the DFGA. Topographic results showed that optimal sets differed among users, which reinforces the need to customize a channel set for each of them. The proposed method computes an optimal subset for each number of channels, allowing the user to select the most suitable set for further BCI sessions.

Brain–Computer Interface web browser for severely disabled people (Martínez-Cagigal et al., 2017).

Víctor Martínez-Cagigal, Javier Gomez-Pilar, Daniel Álvarez, and Roberto Hornero. *IEEE Transactions on Neural Systems and Rehabilitation Engineer-*

ing, vol. 25(8), pp. 1332–1342, 2017. Impact factor in 2017: 3.972, D1 (Q1) in "REHABILITATION", and Q1 in "ENGINEERING, BIOMEDICAL" (JCR-WOS).

Abstract: This paper presents an electroencephalographic (EEG) P300-based brain-computer interface (BCI) Internet browser. The system uses the"odd-ball" row-col paradigm for generating the P300 evoked potentials on the scalp of the user, which are immediately processed and translated into web browser commands. There were previous approaches for controlling a BCI web browser. However, to the best of our knowledge, none of them was focused on an assistive context, failing to test their applications with a suitable number of end users. In addition, all of them were synchronous applications, where it was necessary to introduce a "read-mode" command in order to avoid a continuous command selection. Thus, the aim of this study is twofold: 1) to test our web browser with a population of multiple sclerosis (MS) patients in order to assess the usefulness of our proposal to meet their daily communication needs; and 2) to overcome the aforementioned limitation by adding a threshold that discerns between control and non-control states, allowing the user to calmly read the web page without undesirable selections. The browser was tested with sixteen MS patients and five healthy volunteers. Both quantitative and qualitative metrics were obtained. MS participants reached an average accuracy of 84.14%, whereas 95.75% was achieved by control subjects. Results show that MS patients can successfully control the BCI web browser, improving their personal autonomy.

Towards an accessible use of smartphone-based social networks through Brain–Computer Interfaces (Martínez-Cagigal et al., 2019a).

Víctor Martínez-Cagigal, Eduardo Santamaría-Vázquez, Javier Gomez-Pilar, and Roberto Hornero. *Expert Systems with Applications*, vol. 120, pp. 155-166, 2019. Impact factor in 2019: 5.452, D1 (Q1) in "OPERATIONS RESEARCH & MANAGEMENT SCIENCE", Q1 in "COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE", and Q1 in "ENGINEERING, ELECTRICAL & ELECTRONIC" (JCR-WOS).

Abstract: This study presents an asynchronous P300-based Brain–Computer Interface (BCI) system for controlling social networking features of a smartphone. There are very few BCI studies based on these mobile devices and, to the best of our knowledge, none of them supports networking applications or are focused on an assistive context, failing to test their systems with motor-disabled users. Therefore, the aim of the present study is twofold: (i) to design and develop an asynchronous P300-based BCI system that allows users to control Twitter and

Telegram in an Android device; and (ii) to test the usefulness of the developed system with a motor-disabled population in order to meet their daily communication needs. Row-col paradigm (RCP) is used in order to elicitate the P300 potentials in the scalp of the user, which are immediately processed for decoding the user's intentions. The expert system integrates a decision-making stage that analyzes the attention of the user in real-time, providing a comprehensive and asynchronous control. These intentions are then translated into application commands and sent via Bluetooth to the mobile device, which interprets them and provides visual feedback to the user. During the assessment, both qualitative and quantitative metrics were obtained, and a comparison among other state-of-the-art studies was performed as well. The system was tested with 10 healthy control subjects and 18 motor-disabled subjects, reaching average online accuracies of 92.3% and 80.6%, respectively. Results suggest that the system allows users to successfully control two socializing features of a smartphone, bridging the accessibility gap in these trending devices. Our proposal could become a useful tool within households, rehabilitation centers or even companies, opening up new ways to support the integration of motor-disabled people, and making an impact in their quality of life by improving personal autonomy and self-dependence.

1.2 Context: biomedical engineering, biomedical signal processing and evolutionary computation

Biomedical engineering is the branch of science that applies engineering principles to understand, control and modify biologic systems (Bronzino and Peterson, 2014). It covers a wide range of clinical, industrial and academic activities, including both experimental and theoretical research. One of the greatest benefits of biomedical engineering is the identification and resolution of issues and needs of our healthcare system using technology and systems methodology (Bronzino and Peterson, 2014). However, due to its interdisciplinary character, it is unlikely that a single person could be expertized in the entire field, being essential the interplay among experts of more specific scopes. For instance, specific fields such as neuroengineering, rehabilitation engineering and human performance engineering are also closely related to the design and development of BCI systems.

Physiological signal processing has become the key to understand biological systems from an engineering perspective. The different physiologic systems in hu-

man body (e.g., cardiovascular, endocrine, vision, auditory, gastrointestinal, respiratory, nervous) produce a great number of biomedical signals that reflect their behavior over time (Bronzino and Peterson, 2014). By analyzing these signals, the biomedical engineer may monitor the state of the user, detect meaningful changes in these systems or identify pathological conditions. Unfortunately, this analysis cannot usually be performed at first sight, but a processing step is essential to extract and interpret hidden information. Understanding signals as functions that convey information about the behavior or attributes of some phenomena (Priemer, 1990), signal processing aims at extracting features that characterize them by applying mathematical and information theory methodologies (Sörnmo and Laguna, 2005).

Evolutionary computation, indeed, covers a great number of methodologies that have been successfully applied to such purposes. The term, which has become increasingly important in the last decades, represents a family of optimization algorithms inspired by biological evolution (Eiben and Smith, 2003). All of them are based on the reproduction, variation, competition and natural selection of individuals in a population, which share a common objective. In many of these algorithms, a collective and self-organized behavior emerges from individuals that a priori are only able to follow simple rules, known as swarm intelligence (Yang, 2014). In other words, stigmergy arises while the population converges into a common objective, usually referred to as meta-heuristic. While heuristics rely on problem-specific strategies, meta-heuristics generalize them to problem-independent frameworks (Bozorg-Haddad et al., 2017; Yang, 2014). As nature does, the power of meta-heuristics rely on the most basic problem solving technique – that of trial-and-error (Eiben and Smith, 2003). This apparent simplicity, however, does not hinder their efficacy in solving complex optimization problems, which has led to their successful application in industrial fields, such as computer vision, soft computing, signal and image processing, scheduling, and aerospace engineering, among others (Coello and Lamont, 2004).

In this Doctoral Thesis, the aforementioned fields are applied in physiological signals derived from the brain. An overview of the cerebrum anatomy is thus essential to understand the different brain activity measurements, which are further used to implement BCI systems. A detailed description of this topic is introduced in the next section.

1.3 The human brain

The human brain is the central organ of the nervous system, responsible of processing and coordinating the information received by the senses, making decisions and controlling most of the activities of the body (Standring, 2016). Together with the spinal cord, they constitute the central nervous system (CNS). In a nutshell, the main function of the CNS is to respond to events that occur in either the external world or the body by eliciting neuromuscular or hormonal outputs that serve the needs of our organism (Wolpaw and Wolpaw, 2012b). BCI systems essentially employ these outputs to replace or improve CNS functions by generating artificial signals. Therefore, human brain anatomy, functions and the measurement of its activity are briefly introduced below.

1.3.1 Overview of cerebrum anatomy

Brain (i.e., encephalon) is composed by the cerebrum, diencephalon, brainstem and cerebellum. The former is in turn encompassed by the cerebral cortex and the subcortical areas, which include the hippocampus, basal ganglia and amygdala (Standring, 2016). Although subcortical structures, as well as the diencephalon, brainstem and cerebellum, make important contributions to the CNS, lack of note to practical BCIs due to their unapproachable locations by non-invasive techniques (Miller and Hatsopoulos, 2012). The cerebral cortex, by contrast, is relatively easy to access experimentally, and it has become the primary focus in BCI research.

The cortex is the brain's outer layer of neural tissue, divided by the interhemispheric fissure into two paired hemispheres, which are in turn connected beneath the cortex by the corpus callosum (Standring, 2016). It is colloquially known as *gray matter* because of the color that exhibits its large number of neurons, whereas the *white matter* refers to the many nerve fibers that interconnect the different cortical areas and connect them to subcortical structures (Miller and Hatsopoulos, 2012). The cerebral cortex of higher mammals has seemingly evolved from a relatively smooth sheet to a highly convoluted surface, composed by a set of sulci (i.e., ridges) and gyri (i.e., grooves). As shown in the Figure 1.2, these convolutions divide the cortex into four major anatomical lobes: frontal, parietal, occipital and temporal (Miller and Hatsopoulos, 2012). Even though this division have traditionally relied on an anatomical rationale, it has been demonstrated that lobes are specialized in carrying out the bulk of certain basic functions, which will be discussed afterwards.

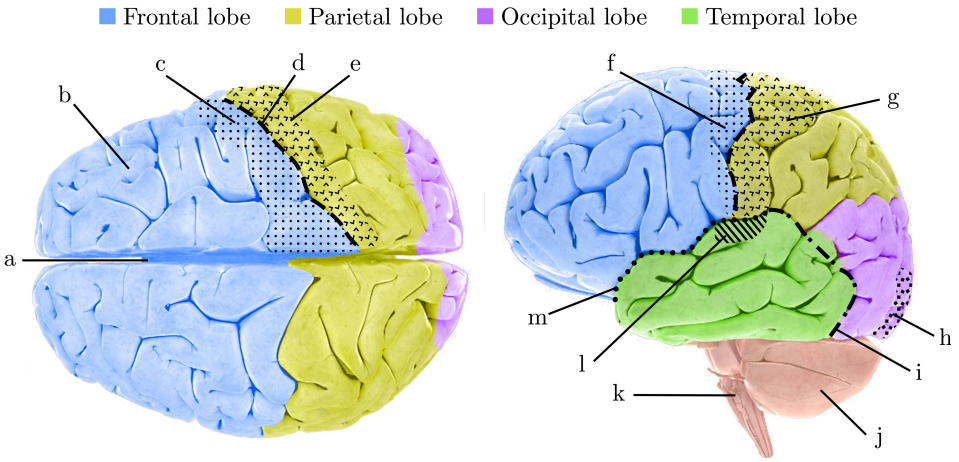


Figure 1.2: Primary anatomical areas of the human cerebral cortex from above (left) and lateral (right) views. The major four lobes are shaded by different colors: blue, frontal lobe; yellow, parietal lobe; purple, occipital lobe; and green, temporal lobe. Anatomical tags corresponds to: (a) interhemispheric fissure, (b) prefrontal association cortex, (c) precentral gyrus, (d) central sulcus, (e) postcentral gyrus, (f) primary motor cortex, (g) primary somatic sensory cortex, (h) primary visual cortex, (i) preoccipital notch, (j) cerebellum, (k) brainstem, (l) primary auditory cortex, and (m) lateral sulcus. Adapted from Wolpaw and Wolpaw (2012a) and Society for Neuroscience (2017).

1.3.2 Brain functional specialization

Since the beginning of the 19th century, researchers have tried to map different body functions over the cerebral cortex (Miller and Hatsopoulos, 2012). The theory of modularity, extended from the outdated phrenology, suggested that the brain has highly specialized regions which are domain specific for different cognitive functions (Fodor, 1983). By contrast, the theory of distributed processing suggested that brain is immensely interconnected and the information is processed in a distributed manner (McIntosh, 1999). However, more recent studies based on graph theory suggest that the functional behavior of the brain is a combination of both (Bullmore and Sporns, 2009; Stam, 2004). Although complex cognitive processes involve the interaction between different parts of the brain, there is a general consensus that particular areas are more specialized than others (Miller and Hatsopoulos, 2012). In BCI research, focusing the activity measurements on specific cortical areas, which should be related to the employed paradigm, is essential to extract suitable features. Among them, the most common key points are the following:

Primary motor cortex (M1). Located in the frontal lobe (Figure 1.2, f), M1

area is closely related to movement control (Miller and Hatsopoulos, 2012). As could be noticed in the representation of Penfield's homunculus (Figure 1.3), particular regions of M1 are intended to control particular body areas. It is noteworthy that the cortical homunculus (i.e., *little man* in latin) is disproportionately depicted over the M1. This is because controlling complex body parts (e.g., hands, facial muscles) requires larger cortical areas than managing simpler ones (e.g., elbow, knee).

Primary somatosensory cortex (S1). S1 is located in the parietal lobe, on the other side of the central sulcus (Figure 1.2, g). This area is also important for movement because it is related to sensations such as temperature, pain, touch or proprioception (i.e., sense of limb position and movement) (Miller and Hatsopoulos, 2012). Many of the axons of S1 neurons connect this region with the thalamus and the spinal cord, for further processing of sensory modalities and spinal reflexes (Miller and Hatsopoulos, 2012).

Prefrontal cortex (PFC). Located in the posterior part of the frontal lobe (Figure 1.2, b), PFC appears to have an important role in high-level executive functions related to movement, planning complex cognitive and social behavior, personality expression, short-term memory and decision making (Miller et al., 2002; Wolpaw and Wolpaw, 2012a). The most dramatic example of loss of these functions due to prefrontal damage is likely the case of Phineas Gage, an American railroad foreman. His frontal lobe was almost completely destroyed after a one-inch diameter iron rod went through his head. Although he survived the accident and could live normally in many aspects, his personality drastically changed, becoming impulsive, unreliable and unable to carry out his future plans (Miller and Hatsopoulos, 2012).

Primary visual cortex (V1). V1 is located in the anterior part of the occipital lobe (Figure 1.2, h). Visual information that comes from the retina goes through the lateral geniculate nucleus (i.e., a relay center in the thalamus) and then reaches the visual cortex. V1 is specialized in processing sensory inputs from the contralateral visual field (Standring, 2016). Hence, superior retinal quadrants (i.e., inferior visual field) are connected with the upper part of the V1, whereas inferior retinal quadrants (i.e., superior visual field) are connected with the bottom part of the V1, and peripheral areas of the retina are connected with the most anterior parts of V1. It is noteworthy that visual information is not processed in a spatial way, but as edge detection (Barten, 1999). For an image containing half side black and half side white,

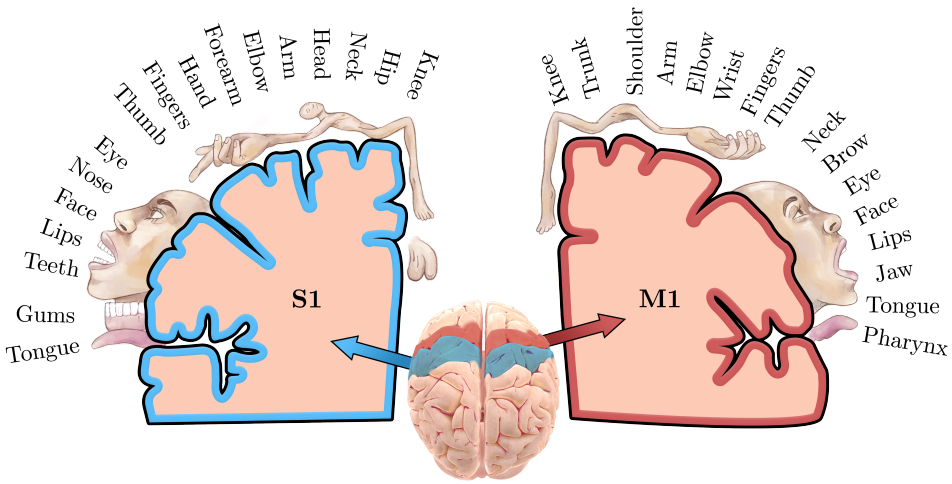


Figure 1.3: Cross section of sensory and motor homunculi over the primary somato sensory cortex (S1, left) and the primary motor cortex (M1, right), respectively. Adapted from Wolpaw and Wolpaw (2012a) and Society for Neuroscience (2017).

the dividing line will produce the strongest contrast due to lateral inhibition. Therefore, most neurons will process the contrast line, while few neurons will encode the brightness information (Barten, 1999).

Primary auditory cortex (A1). Located in the superior and medial parts of the temporal lobe (Figure 1.2, l), A1 is specialized in processing auditory information, being involved in higher functions such as hearing, language switching or identifying the location of a sound in space (Pickles, 2013). Curiously, A1 neurons are arranged according to the frequency to which they react best. In physiology, this frequency arrangement is known as tonotopic map, whose function, although still unknown, appears to reflect that cochlea is also organized according to sound frequencies (Pickles, 2013).

1.3.3 Brain activity measurement

The brain activity that occurs in the aforementioned cortical areas reflects our interaction with the outside world and our responses to external stimuli. When designing a BCI system, it is essential to consider the main advantages and disadvantages of the different brain activity measurement techniques (Wolpaw and Wolpaw, 2012b). From a microscopic point of view, brain activity is generated by an exchange of neurotransmitters between neurons. This chemical phenomenon

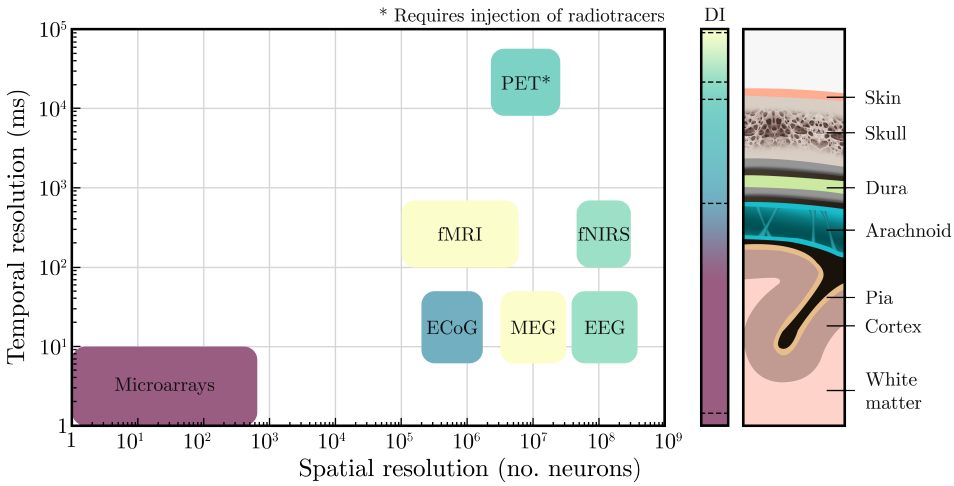


Figure 1.4: Temporal and spatial resolution of different brain activity measurement techniques. Degree of invasiveness (DI) is also represented as a color gradient, according to the brain layers (right). ECoG: electrocorticography, MEG: magnetoencephalography, EEG: electroencephalography, fMRI: functional magnetic resonance imaging, fNIRS: functional near-infrared spectroscopy, PET: positron emission tomography. Adapted from Wolpaw and Wolpaw (2012a).

causes an imbalance in ion concentrations between the inside of a specific neuron and the extracellular fluid, leading to a quick depolarization that propagates across the axon. In practice, the membrane potential of the neuron rises and falls, creating an action potential. When the electrical current reaches the end of the axon, neurotransmitters are released again from the synaptic vesicles, providing a cell-to-cell-communication (Standring, 2016).

BCIs make use of different acquisition methodologies to observe the macroscopic effect of these interactions and thus, to monitor users' brain activity. Figure 1.4 depicts the most popular techniques, classified in function of their temporal and spatial resolutions, as well as of their degree of user invasiveness. Based on their underlying principles, brain activity measurements could be divided into those that record electric, magnetic or metabolic activity:

Electric. Perhaps, the most intuitive way to monitor brain activity is to identify electric sources related to the synaptic activity of individual neurons. Depending on the spatial resolution of the electrodes, electric activity can be measured in *micro-*, *meso-* and *macro-* scales of cortical tissue. Microscale fields (i.e., local field potentials), which reflect the activity of tissue volumes of 0.001–1 mm^3 (equiv. individual neurons), could be recorded by microar-

rays or spikes within the brain. Mesoscale fields, which spread throughout tissue volumes of 1–20 mm^3 (equiv. neural modules, macrocolumns), are typically recorded by electrocorticography (ECoG) from the surface of cortex. Macroscale fields, which reflect joint electrical activity across volume tissues of 10^3–10^4 mm^3 (equiv. brain areas and lobes), are recorded by electroencephalography (EEG) from the user's scalp (Nunez, 2012). As shown in the Figure 1.4, all of them present excellent temporal resolutions (i.e., 1–15 ms). Nevertheless, there is a trade-off between spatial resolution and degree of invasiveness: microarrays and ECoG involve placing electrodes within the cortex and over the arachnoid, respectively; whereas EEG is absolutely not invasive.

Magnetic. The same electric currents that generate the electric field also produce a magnetic field that can be measured by magnetoencephalography (MEG). However, due to the low frequency of brain signals, electric and magnetic fields generated by the brain are uncoupled (Srinivasan, 2012). In other words, MEG is concerned on measuring a *quasistatic* magnetic field that is not related with the electric field that reflects the EEG. Notwithstanding the complementary information that MEG provides, the brain magnetic field is very small relative to unavoidable ambient magnetic variations. Therefore, it is required to use a superconducting quantum interference device (SQUID) magnetometer at a very low temperature and in a specially shielded chamber (Srinivasan, 2012). Besides its lack of portability, MEG is non-invasive and provides an excellent temporal resolution (i.e., 1–15 ms) and a spatial resolution comparable to EEG (Figure 1.4).

Metabolic. When a neuron increases the firing rate of action potentials, it uses more energy. This energy is provided in form of glucose and oxygen, increasing the blood blow in the region. Many techniques use blood flow as a marker to indirectly measure brain activity (Wolpaw and Wolpaw, 2012c). Among them, positron emission tomography (PET), functional near-infrared spectroscopy (fNIRS) and functional magnetic resonance imaging (fMRI) stand out. By injecting a radiotracer in the patient's body, PET is able to track blood flow by detecting pairs of gamma rays caused by electron-positron annihilation. Temporal and spatial resolutions are poor, the method is invasive and involves exposure to ionizing radiation (Wolpaw and Wolpaw, 2012c). By contrast, fNIRS and fMRI are able to track blood flow in a non-invasive way, making PET unsuitable for BCI applications where ease

of use is crucial. Both of them rely on measuring the blood-oxygen-level-dependent (BOLD) response. In our bodies, the transportation of oxygen is carried out by hemoglobin (Hb), an iron-containing protein that coexists in two forms: deoxy-Hb (without oxygen) and oxy-Hb (with oxygen). As blood passes through the lungs, each deoxy-Hb picks up four O_2 molecules, becoming oxy-Hb; whereas when blood passes through organs and muscles, oxy-Hb molecules release oxygen and revert back to deoxy-Hb. Whenever a brain region increases its activity, a BOLD response requests more oxy-Hb molecules, increasing the blood flow. Hemoglobin presents different properties that allow BOLD response being tracked: oxy-Hb is light red and nonmagnetic, whereas deoxy-Hb is dark red and slightly magnetic (Wolpaw and Wolpaw, 2012c). In fNIRS, probes emit infrared light at specific wavelengths and measure the absorption when light has passed the underlying tissue. Differences in the absorption spectrum of oxy-Hb and deoxy-Hb makes possible to monitor BOLD responses in different brain regions (Ferrari and Quaresima, 2012). Conversely, fMRI relies on the magnetic properties of the different states of Hb. MRI scans use strong magnets and brief radio-frequency pulses to modify the magnetic moment of nuclear protons. Deoxy-Hb, due to its paramagnetic property, provides a disturbance on the moment additional to the magnetic field created by the magnet, allowing fMRI to measure the BOLD response (Ogawa et al., 1990). Even though the spatial resolution of fMRI is excellent and fNRIS's one is comparable to EEG, the temporal resolution of both techniques is limited by the slow rate of the BOLD response, which lasts up to several seconds (Wolpaw and Wolpaw, 2012c). Moreover, effective fMRI neuroimaging requires many scans to remove external artifacts (i.e., movements, respiration, vessels pulsations) and an expensive and non-portable equipment.

Practical BCIs usually are required to work in real-time, making it essential to have an excellent temporal resolution. For that reason, metabolic-based physiological signals are often relegated to the research field. Among the rest of them, most BCIs prioritize the portability of the system and avoid invasive approaches, preferring EEG over other measurement techniques.

1.4 Brain–Computer Interfaces

Since the invention of the EEG by the professor Hans Berger in 1929, scientists have speculated about the possibility of moving or communicating through brain signals rather than muscles (Berger, 1929). Nevertheless, it took until 1973 to publish the first BCI in the literature. The system, developed by the professor Jacques Vidal, was able to detect the direction of the eye gaze by processing visual evoked potentials with the objective of controlling a two-dimensional cursor (Vidal, 1973b, 1977). Even though there is still a lot to be done, BCIs have drastically evolved since then to the point of being successfully applied with motor-disabled people to improve their quality of life; or even of becoming commercially interesting in many non-clinical applications.

1.4.1 A closed loop

BCIs provide a communication pathway between the brain and the environment, making it possible to control external devices or applications by using the user's brain activity. More specifically, Wolpaw and Wolpaw (2012a) define this emerging technology as:

> *"(A BCI is) a system that measures CNS activity and converts it into artificial output that replaces, restores, enhances, supplements, or improves natural CNS output and thereby changes the ongoing interactions between the CNS and its external or internal environment."*

This definition not only describes what a BCI is, but also the five most straightforward applications. In practice, BCIs might: (1) replace natural outputs that have been lost due to injuries or diseases (e.g., speech synthesizer for people that cannot longer speak); (2) restore weak responses (e.g., muscle stimulation via implanted electrodes for spinal cord injuries); (3) enhance CNS outputs (e.g., sound alert for warning against somnolence when driving); (4) supplement neural outputs (e.g., control of an extra robotic hand); and (5) improve CNS responses (e.g., rehabilitation of muscles via sensorimotor brain signals in stroke patients) (Brunner et al., 2015; Wolpaw and Wolpaw, 2012a). Independently of their final purpose, BCI systems share a closed loop structure. The three main stages that compose them, depicted in the Figure 1.5, are detailed below:

1) **Signal acquisition.** The first stage involves the recording of brain activity. As discussed in section 1.3.3, there are several methods that monitor electrical, magnetical or metabolic activity from the brain at different levels of

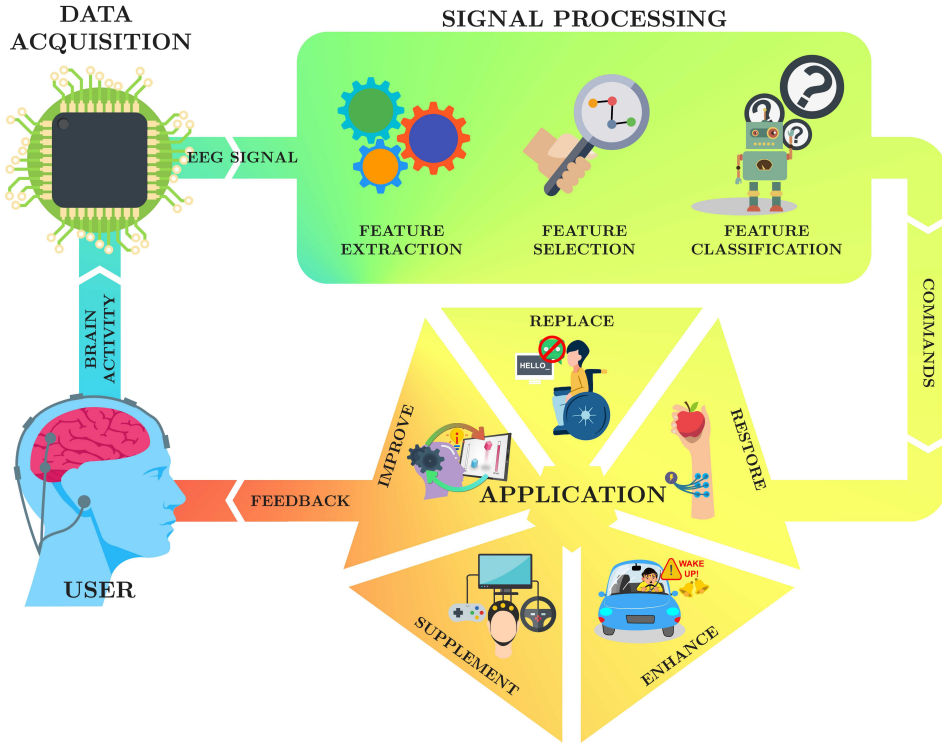


Figure 1.5: Classical structure of a BCI system. Firstly, user's brain activity is acquired using EEG electrodes. Then, user's intentions are decoded by the signal processing stage, which involves feature extraction, selection and classification steps. Finally, the decoded intentions (i.e., commands), are sent to the final application, which provides feedback to the user. As shown, applications are intended to replace, restore, enhance, supplement and/or improve the natural nervous outputs of the user.

invasiveness. Owing to the practicality of the system, most BCI studies in the literature employ the EEG because of its non-invasiveness, low-cost and ease of use.

2) **Signal processing.** Users' intentions are not directly reflected in their brain signals. The second stage is thus intended to process the acquired signals in order to detect measurable changes that are correlated with the users' intentions. Signal processing is in turn composed by feature (1) extraction, (2) selection, and (3) classification stages. The former aims to extract several *features* somewhat correlated with the intentions by applying signal processing methodologies. Selection algorithms are then applied to select the most relevant features for preventing the curse of dimensionality

(Blankertz et al., 2011). Finally, feature classification makes use of machine learning algorithms to identify the intentions of the user and convert them into application commands.

3) **Application.** Although signal acquisition and processing stages are usually shared in BCI systems, the third stage depends utterly on the study. The application is intended to receive and interpret the commands for providing a real-time feedback to the user. This feedback is essential to close the loop, allowing users to react by selecting new commands or modulating their own brain activity.

1.4.2 EEG-based systems

From a practical point of view, BCIs for general population should ideally be non-invasive, portable, cheap, reliable, easy to setup, comfortable, and robust against different environments and external artifacts. Attending to these aspects, techniques that are invasive, such as microarrays, ECoG and PET; cumbersome or expensive, such as MEG, fMRI and fNRIS; are often delegated to the research field. Therefore, EEG-based BCIs have been prioritized in the present Doctoral Thesis, because of their non-invasiveness, low-cost and ease of use.

Electroencephalogram is a non-invasive monitoring technique that records electrical activity from the brain by placing a set of electrodes over the user's scalp. Each electrode reflects a space-averaged activity of 100 million to a billion of neurons, whose coordinated firing is strong enough to be recorded from outside the brain (Nunez et al., 1997; Wolpaw and Wolpaw, 2012a). A typical EEG recording involves the use of, at least, three electrodes (ground, reference and channel), a differential amplifier, hardware filtering, an operational amplifier and an analog-to-digital converter (ADC). The ground electrode is connected to the amplifier ground, which is isolated from the power supply, preventing amplifier drifts and favoring the common-mode rejection (Srinivasan, 2012). Note that EEG recordings are bipolar, and thus they measure voltage differences between a channel and a reference. The reference electrode should be ideally located far enough from brain sources, which is usually placed at the ear, mastoid or neck. Electrodes are directly connected to a differential amplifier, intended to amplify the difference between each channel and the reference, and to suppress the common voltage between them (i.e., common-mode). Remaining common voltages are due to unequal impedances between electrodes, which are lessened by applying a conductive gel between the electrodes and the scalp (Ferree et al., 2001). Then, hardware filters

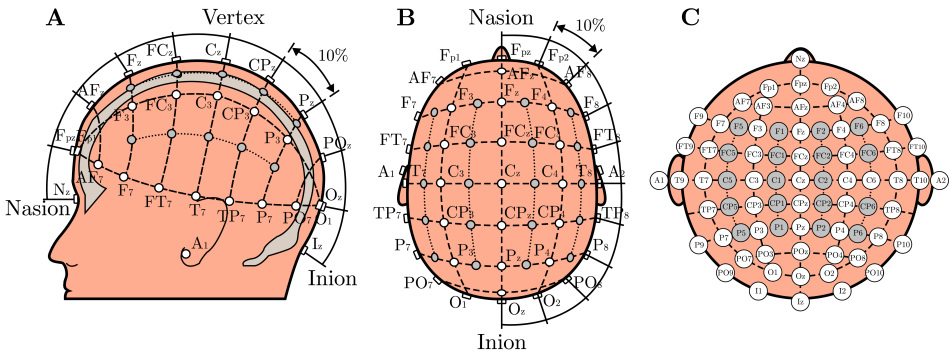


Figure 1.6: Schematic views of the standard International System 10–10 montage for EEG recordings: (A) lateral, (B) top, and (C) scalp projection. In (A) and (B), gray electrodes are not labeled for clarification purposes. This standard uses proportional distances of 10% between the nasion and the inion for locating the landmarks.

typically remove frequencies below 0.5 Hz and over 200 Hz, although it depends on the implementation. Signals are finally amplified about $2 \cdot 10^4$ times, sampled and digitized. Channel positions are not arbitrary, but are specified in different standards, such as International Systems 10–20, 10–10 or 10–5 (Srinivasan, 2012). In the present Doctoral Thesis, the standard 10–10 is applied, which establishes landmarks following proportional distances of 10% between the nasion and the inion (Chatrian et al., 1985), as depicted in the Figure 1.6.

The high sampling rate of the ADC (commonly, $f_s \geq 250$ Hz) guarantees a suitable temporal resolution (i.e., 1–15 ms), good enough to capture electrical aftermaths of neurological processes in real time. However, the spatial resolution of the EEG is limited due to the distance between cortical sources and the scalp. As shown in the Figure 1.7, the cortical sources reflected by the EEG are caused by large assemblies of pyramidal neurons that are arranged perpendicularly to the cortex's surface (Nunez et al., 1997). From a macroscopic point of view, each source may be simplified as a dipole, making EEG more sensitive to correlated dipoles on gyri than on sulci. In fact, sulci folds provoke the cancellation of their electrical sources, which cannot be recorded by EEG channels (Srinivasan, 2012). Moreover, the individual contribution of a single source is hindered by the set of tissues that must go through. This tough journey is characterized by two phenomena: (1) source effect, and (2) volume conduction. The former involves the attenuation, distortion and noise contamination of a source when it passes through the cortex, meninges and the skull (Wang, 2010). Volume conduction is concerned with the spread of electrical activity when it encounters a different surface, such as the scalp

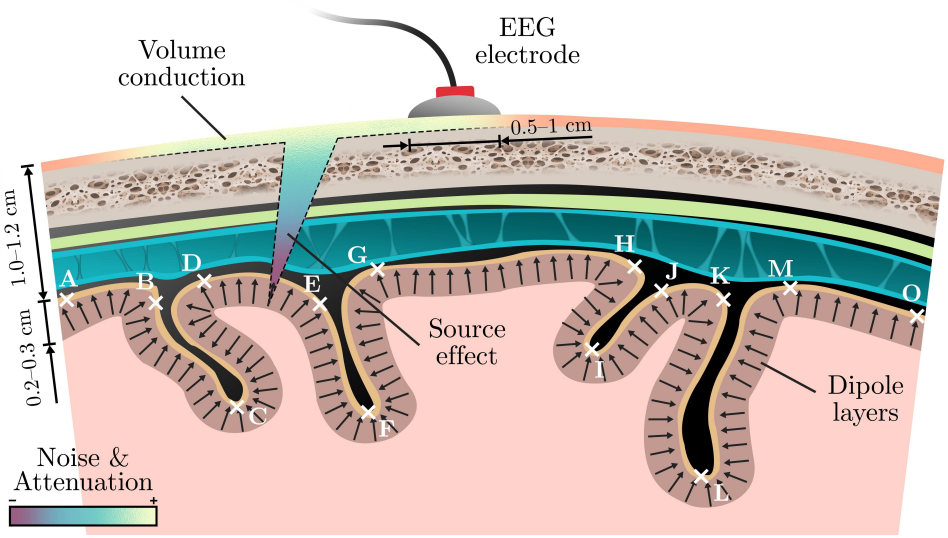


Figure 1.7: Anatomical illustration of cortical sources perceived by the EEG. Pyramidal cells are depicted as dipole layers that are perpendicular to the cortical sheet. EEG is more sensitive to correlated dipoles in gyri perpendicular to the electrode (e.g., G–H); less sensitive to correlated dipoles that are not completely perpendicular (e.g., D–E); and insensitive for such dipoles that cancellate themselves, such as those found in sulci (e.g., B–C–D, E–F–G, H–I–J, K–L–M). Source effect and volume conduction phenomena are also depicted.

(Nunez et al., 1997). Even though the laws that govern the volume conduction effect are well-known, their application to EEG is extremely complex because of the time dependence (Nunez, 2012). It is necessary to take into consideration that both effects are inherent to EEG recordings.

The attenuation and distortion that neural sources suffer when traveling from the cortex to the scalp makes EEG sensitive only to the coordinated activity of billions of neurons at the same time. For this reason, EEG is often interpreted as a rhythmic activity that reflects neural oscillations (Cohen, 2014). Although the frequency spectrum is not limited, the oscillations that are thought to be associated with cognitive processes are comprised in frequencies between 2–150 Hz (Cohen, 2014). As shown in the Figure 1.8, EEG signals may be further decomposed into five main frequency rhythms or bands:

1) **Delta (δ, 1–4 Hz).** The slowest frequency band, characterized by high-amplitude waves, prevails frontally in adults in deep sleep. It is also seen in awake babies, but spatially located at occipital electrodes. An excessive and generalized delta activity is abnormal, commonly associated with subcortical,

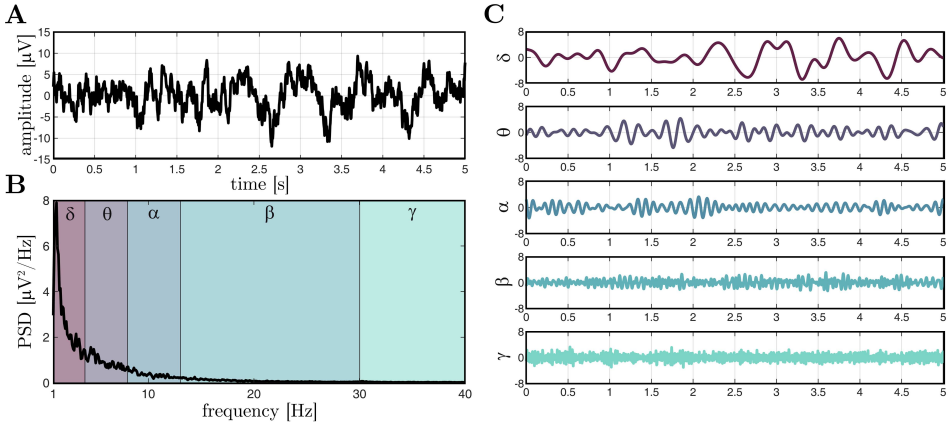


Figure 1.8: Neural oscillations of electroencephalographic (EEG) signals of a healthy awake subject. (A) EEG segment of 5 seconds from channel Pz; (B) power spectral density (PSD) of the EEG signal; (C) decomposition of the EEG segment in the five main bands: δ (1–4 Hz), θ (4–8 Hz), α (8–13 Hz), β (13–30), and γ (30–100 Hz).

deep midline or diffuse lesions, as well as with metabolic encephalopathy hydrocephalus (Tatum IV, 2014).

2) **Theta (θ, 4–8 Hz).** Predominant in frontal or frontocentral regions, theta activity is common in young adults. These waves are commonly enhanced by drowsiness or hyperventilation; and are associated with meditative, emotional and creative states (Tatum IV, 2014). Intermittent bursts of generalized theta activity are often abnormal, which could indicate mild diffuse encephalopathy or hydrocephalus (Tatum IV, 2014).

3) **Alpha (α, 8–13 Hz).** Also known as the *posterior basic rhythm*, the alpha band was the first reported rhythmic activity by Hans Berger (Berger, 1929). These waves are most prominent in bilateral posterior head regions during relaxed wakefulness. For most adults, alpha rhythms increase their amplitude immediately after eye closing, and attenuate in eye opening or mental exertion. Unilateral failure may reflect an ipsilateral abnormality (Tatum IV, 2014). Furthermore, alpha band is also referred as mu rhythm (μ, 8–13 Hz) when its topography is mainly located over the sensorimotor cortex (i.e., S1 and M1 regions). The mu rhythm is a decrease of contralateral alpha activity associated with limb movements, or even the imagination of these movements (Reilly, 2013; Wolpaw et al., 2002). It may be asymmetrical or only seen at one side, which do not necessarily disclose a brain lesion. By contrast, abnormal behavior is reflected if mu rhythm is persistent or not

reactive to motor execution (Tatum IV, 2014).

4) **Beta (β, 13–30 Hz).** Frontal low-amplitude beta rhythms are closely linked to anxiety, active concentration, drowsiness or light sleep. An excessive activation in this band is often associated with the consumption of benzodiazepines, barbiturates and chloral hydrates; as well as with pathologies such as the *Dup15* syndrome (Tatum IV, 2014).

5) **Gamma (γ, 30–100 Hz).** Less studied, gamma rhythms are thought to be related with high cognitive processing. These waves arise during cross-modal sensory processing (perception that involves two or more senses at the same time) and short-term memory matching (Kisley and Cornwell, 2006). Recent studies also associate gamma activity with abrupt interactions between excitatory and inhibitory neurotransmitter concentrations (Fuchs et al., 2007).

1.4.3 Control signals

The most important role of a BCI system is to accurately detect the intentions of the user in real-time, which is not a straightforward process. As aforementioned, complex cognitive processes such as thinking or decision making involve the interaction between different parts of the brain (section 1.3.2). Furthermore, the spatial resolution of EEG and its inability to record neither non-perpendicular pyramidal neurons or deep activity would hinder the focusing on certain brain areas (section 1.4.2). Under this rationale, users' intentions are not directly reflected on their raw EEG signals, making paradigms and signal processing steps essential to identify and convert them into application commands. Actually, EEG-based BCI systems rely on the processing of measurable changes related to cognitive tasks, known as *control signals*.

Control signals may be divided in function of the required time to master them. Exogenous signals rely on processing natural responses of our brains to external stimuli; thus, do not require any training, but a short calibration (Kleber and Birbaumer, 2005; Nicolas-Alonso and Gomez-Gil, 2012). By contrast, endogenous signals are based on users' ability to self-regulate their brain activity through neurofeedback (NF) training (Nicolas-Alonso and Gomez-Gil, 2012). Based on operant conditioning, NF involves a real-time presentation of certain EEG parameters to the user, together with positive reinforcement for favoring the learning of the desired regulation (Kleber and Birbaumer, 2005). In practice, users eventually find their own strategies to self-regulate their brain rhythms in order to complete a

certain task or, in this case, to generate measurable changes in their EEG signals. Unfortunately, the control of these endogenous signals is time-consuming, lasting from hours to weeks, depending on the user (Nicolas-Alonso and Gomez-Gil, 2012). Notwithstanding the clear distinction between both types of control signals in function of the training time, this taxonomy may be confusing because there are control signals that encompass both exogenous and endogenous components (Donchin et al., 1978).

Slow cortical potentials

Slow cortical potentials (SCP) are a family of slow endogenous voltage shifts that are time-locked and phase-locked to certain sensorimotor activities (i.e., occur at specific times before, during and after the events) (Allison Brendan Z., 2012). These events are triggered by executing or imaging movements, or by achieving cortical activation through cognitive tasks (e.g., mental arithmetic, concentration) (Allison Brendan Z., 2012; Nicolas-Alonso and Gomez-Gil, 2012; Wolpaw et al., 2002). SCPs are characterized by a slow negative wave, which starts 1.5−2 s prior to the movement onset and reaches its peak negativity at the onset; followed by a positive rebound that usually lasts 1−1.5 s (Guger et al., 2014). Their characteristic slowness makes them appear in low delta bands (<1 Hz). In the literature, SCPs are often referred as movement-related cortical potentials (MRCP), since they are thought to reflect neural activation in preparation for action (Allison Brendan Z., 2012; Guger et al., 2014; Jahanshahi and Hallett, 2003).

MRCP (i.e., SCP) waveforms differ depending on the NF paradigm and their source localization across the cortical surface. Notwithstanding the numerous components that make the MRCPs up, whose nomenclature and taxonomy vary among authors (Allison Brendan Z., 2012; Farina et al., 2013; Jahanshahi and Hallett, 2003), recent BCI studies distinguish between the *bereitschaftspotential* (BP) and the *contingent negative variation* (CNV) (Guger et al., 2014). Both are subcomponents of the MRCPs, but generated using different NF paradigms: self-paced (i.e., asynchronous) for BP; and cue-based (i.e., synchronous) for CNV. The BP, also referred as readiness potential (RP), consists of a negative voltage deflection that usually begins 0.5−1.5 s before a volitional movement (Jahanshahi and Hallett, 2003). Although it was originally believed that BP sources lain in subcortical structures, such as basal ganglia and thalamus, more recent studies suggest that sensorimotor areas (M1 and S1) are probably generators of BPs (Ikeda et al., 1992; Toma et al., 2002). CNVs (or anticipation-related potentials) , by contrast, are produced when the user anticipates motor or cognitive tasks (Guger

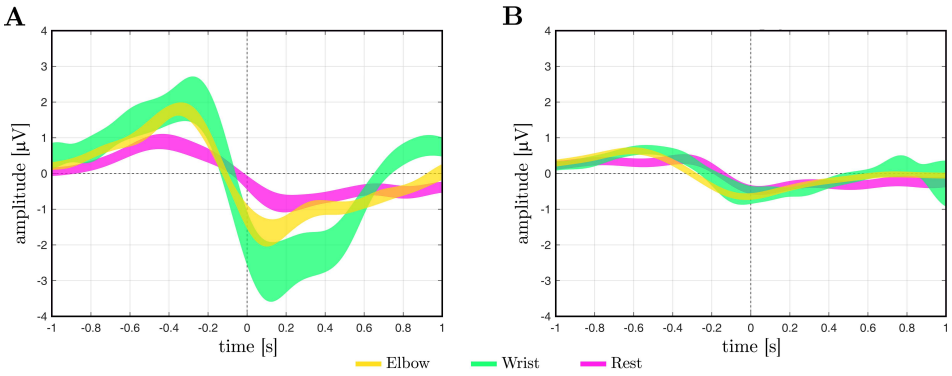


Figure 1.9: Grand averaged MRCPs of 15 subjects performing motor movement tasks with their elbow and wrist. Shaded areas represent the 95% confidence interval across trials. (A) Motor execution response over Cz. (B) Motor imagery response over FCz. MRCPs were computed using a public EEG dataset of upper limb movements (Ofner et al., 2017).

et al., 2014). These waves are negative EEG drifts elicited 0.2−0.5 after displaying a cue that warns an imperative stimulus will occur several seconds later. Since a CNV reflects preparation for response, it is highly affected by motivational and task-specific aspects. Besides being distributed over the regions that are directly involved in the motor action, CNVs are mainly generated in frontal areas, such as the PFC (Allison Brendan Z., 2012; Lu et al., 2012).

Regardless of their main subcomponent, a successful elicitation involves previous training, which usually requires repeated sessions over weeks or months. Moreover, SCPs only offer dichotomous decisions: cortical inhibition or activation. SCP-based BCIs typically dissociate the problem-specific set of commands in a tree-based selection framework, allowing users to select any command by sequentially performing binary selections. This scheme ends up hindering the speed and ease of use of the final application, reaching low transmission rates and making the BCI prone to errors (Allison Brendan Z., 2012; Nicolas-Alonso and Gomez-Gil, 2012). For these reasons, SCPs in BCI systems have been gently relegated and substituted by more reliable control signals. Further SCP applications might be focused on supplementing hybrid BCIs with a complementary measure to anticipate users' movements or to reach an asynchronous control, enhancing the restoring of muscle functions for rehabilitation purposes, or performing basic research or neuromarketing, among others (Allison Brendan Z., 2012).

Sensorimotor rhythms

Sensorimotor rhythms (SMR) are oscillations in mu (8–13 Hz), beta (18–30 Hz), and gamma (30–100 Hz) bands recorded over the sensorimotor cortex (i.e., S1 and M1 areas). SMRs are modulated by endogenous motor activity, leading to EEG variations that are time-locked to the event, but not phase-locked (Pfurtscheller and Lopes da Silva, 1999). In other words, the power of SMRs may vary in association with actual or imagined sensorimotor events, experimenting either a (1) decrease, known as *event-related desynchronization* (ERD); or an (2) increase, known as *event-related synchronization* (ERS) (Pfurtscheller and Lopes da Silva, 1999; Pfurtscheller and McFarland, 2012).

In particular, mu band SMRs exhibit two different behaviors. For lower frequencies (8-10 Hz), ERDs occur for any kind of motor event across the entire somatosensory region, which suggests a reflection of motor preparation or attentional processes. On the other hand, ERDs produced in higher mu frequencies (10-13 Hz) are topographically and functionally restricted over certain task-specific areas (Pfurtscheller and McFarland, 2012). Note that the cortical representation is contralateral to the movement. These localized mu ERDs are usually accompanied by simultaneous ERSs in neighbor cortical areas, reflecting a phenomenon known as *focal-ERD/surround-ERS*. Is it though that this effect may depict a mechanism that joins efforts on an specific sensorimotor subsystem by inhibiting other non-related cortical areas (Pfurtscheller and McFarland, 2012). Therefore, ERD/ERS variations reflect local interactions between neurons, allowing accentuations/inhibitions of certain oscillations (Pfurtscheller and Lopes da Silva, 1999). Beta SMRs also exhibit ERDs in response to motor events or somatosensory stimulation, but usually followed by a short ERS after movement, known as *beta rebound*.

Owing to the similarity between ERD/ERS variations caused by actual and imagined motor events, SMR-based systems have gained a special interest in the BCI literature. Through NF training, some users can learn to self-regulate their own SMRs, even for subjects whose actual limb control is hindered or non-existent (Pfurtscheller and McFarland, 2012). SMR variations are achieved by means of motor imagery, repeatedly emphasizing kinesthetic experiences rather than visual representations of movements (Nicolas-Alonso and Gomez-Gil, 2012). Eventually, the control become automatic and the imagery turns out to be less important. Nevertheless, current SMR-based BCIs are only able to detect spatially well separated classes, usually two (e.g. left and right hands); although some studies have attempted to control up to four (left and right hands, foot and tongue) (Ofner et al.,

2017; Tangermann et al., 2012). Thus, the dissociation of application commands into a tree-based selection framework is also present in SMR-based BCIs for communication and control, due to the greatly limited number of discriminative classes. In therapeutic applications where an accurate control of ERD/ERS events is not sought, but the reinforcement/reorganizing of neural pathways through NF training, SMR-based BCIs are essential to provide continuous feedback (Pfurtscheller and McFarland, 2012).

Figure 1.10 depicts SMRs variations elicited by a motor imagery task, where the subject was requested to imagine left- and right-hand movements for 4 s after a cue was displayed on a screen. As shown in (A), contralateral ERDs are mainly present 1–4 s after the cue, followed by small ERS rebounds. The spectral analysis (B) over this particular temporal window also displays a clear ERD, reducing the power of the contralateral imagery task in the mu band. Furthermore, a time-frequency analysis (C) of the signed-r^2 statistic (i.e., coefficient of determination) was performed to exhibit the difference between ERD/ERS of both tasks. As can be noticed, this subject reached a higher skilled control of left-hand motor imagery. The higher variation of the rhythm when the left-hand task was performed would lead to an easier discrimination in the classification stage.

Steady-state visual evoked potentials

An event-related potential (ERP) is a pattern of voltage variations produced in response to an external event (e.g., visual, auditory or tactile) (Sellers Eric W., 2012). Particularly, when the ERP is time-locked to a *visual* event, such as a brief flash, appearance of sudden images or abrupt color changes, it is known as visual evoked potential (VEP). VEPs are generated in or near the V1 and are usually more prominent over occipital areas (Allison Brendan Z., 2012). However, single-trial ERPs are masked by EEG background activity which is unrelated to the event, making them impossible to identify in the raw EEG with the naked eye. Averaging across trials allows for extracting ERP patterns, making background activity cancel itself and strengthen the time-locked response (Luck, 2014).

ERP voltage patterns utterly depends on the type of stimulus, as well as on the task or paradigm. For instance, VEP main components include N75, P100 and N135, which usually occur 70, 100, and 135 ms after the stimulus onset, respectively (Sellers Eric W., 2012). If the stimuli are presented in a slow rate, the evoked responses are called *transient VEPs*; i.e., the pattern occurs once and then ends. Nevertheless, if the stimuli is a rapid repetitive visual stimulation, the evoked responses overlap themselves, producing a constant oscillation at the same

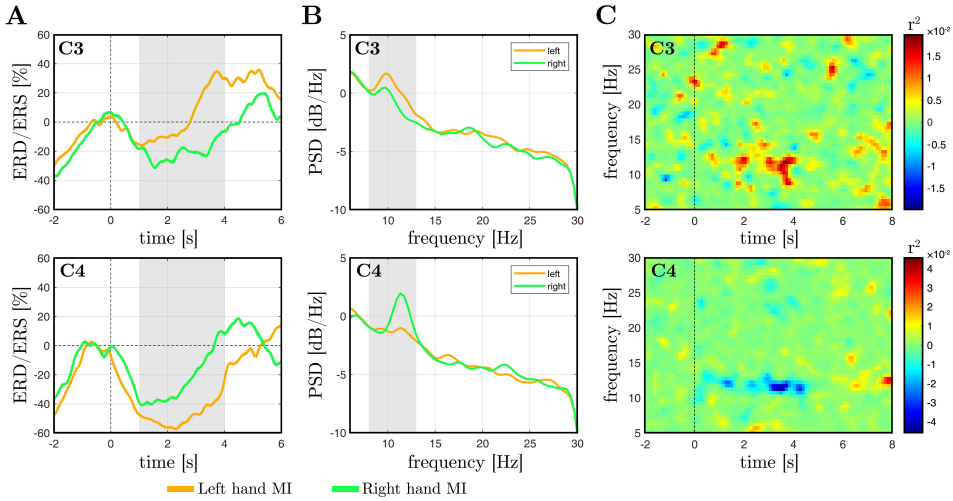


Figure 1.10: Differences between SMRs of left and right hands during a motor imagery (MI) task, where the cue was displayed at $t = 0$ s. (A) Averaged ERD/ERS (%) variations in mu band (8–13 Hz) over C3 and C4. (B) Averaged PSD variations in temporal epochs between 1–4 s over C3 and C4. (C) Spectrograms of signed-r^2 statistics between both classes over C3 and C4 (positive values reflect higher amplitudes of the left class). These SMRs were computed using B subject calibration trials of the public IV BCI Competition dataset 1 (Tangermann et al., 2012).

rate, known as *steady-state VEP* or SSVEP (Luck, 2014; Sellers Eric W., 2012). The Figure 1.11(A) depicts how a simulated transient VEP becomes an SSVEP as the stimulus rate increases.

SSVEPs are easily identifiable by simple spectral analysis, assuring high classification accuracies in SSVEP-based BCIs without prior training, due to their exogenous nature. As shown in Figure 1.11(B), not only SSVEPs are clearly reflected in the PSD spectrum, but also their harmonics. The most common SSVEP-based BCI setup involves creating an arrangement of commands or characters that flicker at different frequencies. When the user pays attention to a specific command, an SSVEP appears in the EEG spectrum and can be detected by the system. Even though this setup allows selecting a high number of different classes, it should be noted that most reliable flickering rates belong to the low beta band (i.e., 13–19 Hz), which increments visual fatigue and maximizes the risk of photosensitive epileptic seizures (Pastor et al., 2003). Moreover, the number of commands is slightly limited by the vertical standardized refreshing rate of current LCD screens (Volosyak et al., 2009).

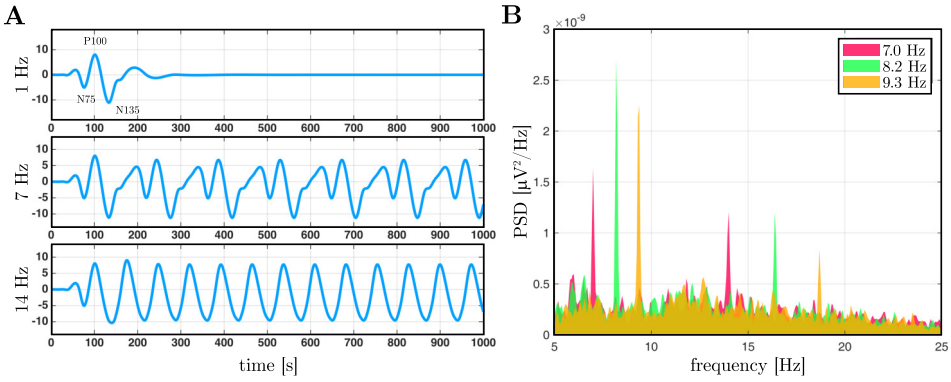


Figure 1.11: (A) Simulated VEP responses at different stimulus rates: 1, 7 and 14 Hz. Note how the transient VEP becomes an SSVEP as the stimulus rate increases. (B) PSDs of real SSVEPs produced as responses to flashes that flickered at 7, 8.2 and 9.3 Hz over the occipital cortex (channel Oz). These SSVEPs were computed using Subject 1 trials of the public AVI SSVEP database (Vilic, 2013).

P300 evoked potentials

The usefulness of ERPs as BCI control signals is not limited to transient VEPs or SSVEPs, but its discriminative power can be strengthened by using more complex events. Voltage patterns that arise when performing a time-locked average of a response to an event are known as *components*. VEP main components, which may be recorded from the first 150 ms following the stimulus onset, tend to reflect cortical activity in V1. Although they usually vary depending on the modality of the stimulus, these initial components are unconsciously triggered by the subject and thus, they are considered exogenous (Sellers Eric W., 2012). Nonetheless, a visual eliciting event may also cause later components if it is related with a certain task that the subject must perform. In that case, longer-latency patterns reflect higher-level cognitive processing less dependent on the stimulus modality, usually referred as endogenous components (Sellers Eric W., 2012).

P300 evoked potentials (i.e., P3) are likely the most studied endogenous component of VEPs. As its name suggests, P300 potentials are positive deflections that appear in response to infrequent and particularly significant stimuli at about 300 ms after their onset (Wolpaw et al., 2002). The latency actually may vary between 250 to 750 ms, though; depending on the spectral filtering, individual aspects of users or on the difficulty of the task (Picton, 1992). These variations suggest that it is indeed an endogenous component, elicited by the decision that a rare event has occurred, whether conscious or not. P300 waves are usually more

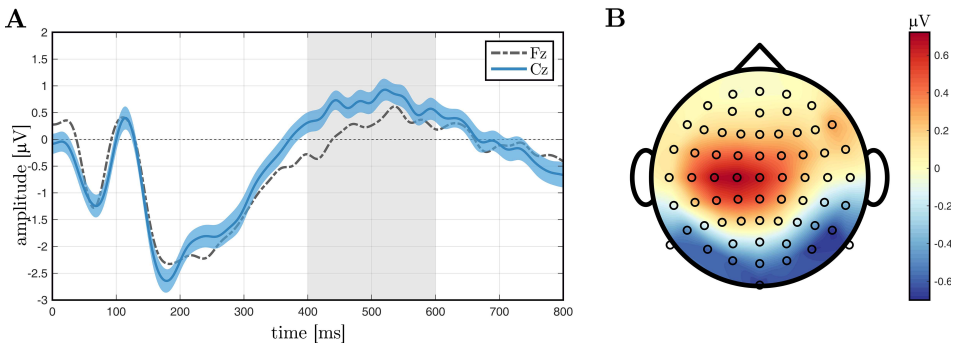


Figure 1.12: Grand averaged P300 evoked potentials across 18 subjects that performed a visual oddball task. (A) Temporal ERP over Fz (dot-dash line) and Cz (solid line). The blue shaded area indicates the 95% confidence interval across trials. (B) Topographical plot of the averaged ERP amplitude between 400 and 600 ms for each channel. These ERPs were computed using a public visual oddball database (Robbins et al., 2018).

prominent at central or parietal scalp positions, and attenuate gently from there (Sellers Eric W., 2012).

The way to generate P300 waves in an ERP is known as *oddball* paradigm. The paradigm is based on presenting infrequent target stimuli, which must be attended; among other background stimuli, which must be ignored. Therefore, the task of the user lie in classifying somehow these two different events. The less-frequent event (i.e., target) elicit a P300, whereas the other one does not (Sellers Eric W., 2012). Note that the lower the probability of the target stimuli, the higher the P300 amplitude. Figure 1.12 depicts a grand averaged P300 across 18 healthy subjects that were performing a visual oddball task (pressing a button when a target image appeared) (Robbins et al., 2018). As shown, the initial temporal components (e.g., N75, P100, N200) are shared with common VEPs (section 1.4.3), but a positive deflection occurs after 380 ms. In this case, the P300 wave is lengthened (until 750 ms) due to the grand average across subjects, since each of them would reflect a different latency. The topographic plot of Figure 1.12(B) displays the mean amplitude between 400–600 ms across the scalp surface. As shown, the P300 potential is more prominent at central regions, whose amplitude declines mainly in occipital areas until it turns into a negative deflection.

P300-based BCIs make use of P300 potentials to detect what command the user wants to select from a set of them. Generally, the system implements variations of the visual oddball paradigm that correlate the P300 response of the subject with the provided commands, identifying the desired command after several stimuli repetitions (section 3.2). In practice, P300-based BCIs do not only process the

P300 wave, but the entire ERP. For that reason, some authors refers that kind of systems as ERP-based BCIs (Santamaría-Vázquez et al., 2019). It is also noteworthy that there is no consensus when referring them as exogenous or endogenous systems, due to the mix of VEP and P300 components. However, it is clear that P300 potentials are generated as natural responses of brains to stimuli and thus, P300-based BCIs do not require any training by the user, just a slight calibration. Moreover, as P300 responses present higher amplitudes when the target stimulus is less frequent, these BCIs favor the inclusion of a high number of different classes (i.e., commands). These aspects make P300-based BCIs so reliable that they can be used by disabled people on a daily basis for communication and/or environmental control purposes (Sellers Eric W., 2012). Owing to the aforementioned advantages, the present Doctoral Thesis is focused on designing, developing and testing P300-based BCIs hereinafter.

1.4.4 Current limitations

The most straightforward applications of BCIs are focused on improving the quality of life of severely disabled people by replacing, restoring, enhancing, supplementing, or improving natural CNS outputs (Wolpaw and Wolpaw, 2012a). However, the level of development of current BCIs is usually not enough to take the leap from laboratories and provide a real-life use of these systems. Indeed, BCI is considered an *orphan technology*, which denotes systems that have been tested in the laboratory, but which do not provide enough incentives to be commercially interesting in their current form (Kübler et al., 2020; Wolpaw Jonathan R., 2012). Therefore, BCI studies intended to provide assistive applications should deal with, at least, the following limitations:

1) **Hardware.** Apart from the inherent limitations of non-invasive BCIs, which were previously discussed in section 1.4.2; EEG equipment should ideally not require conductive gel, be comfortable, easy to setup, fully portable, wireless, function many hours without maintenance, and perform well in real-life environments. Although numerous efforts have been made to develop more and more comfortable and independent equipment, current EEG systems still lack of robust performances in all environments.

2) **Reliability.** Due to the inherent variability of the EEG, BCI performances tend to vary markedly between trials, sessions and individuals. Even though they usually improve with practice, performances never approximate to a

muscle-based control, nor to the required level of reliability for a practical use (Wolpaw Jonathan R., 2012). For MI-based BCIs, reliability is further limited by *BCI illiteracy* (i.e., people who is not able to reach an adequate performance after training sessions) and the small number of classes that can be successfully discriminated.

3) **Validation.** BCI studies often fail to prove their systems with target users (i.e., severely disabled), making it impossible to infer their feasibility in a real context (Martínez-Cagigal et al., 2019a). Instead, most studies only test their systems with healthy users, providing an initial approximation of the expected behavior. Nonetheless, care must be taken when generalizing these results, since it has been widely documented that disabled people usually reach lower performances than control subjects (McCane et al., 2014). Therefore, recruiting a target population must be essential to assess the viability of the proposed system.

4) **Synchrony.** Ideally, a BCI should evolve toward an unsupervised use of the system that provides users with a comprehensive self-paced control. Unfortunately, most studies propose synchronous BCIs, which implies that the system is constantly translating EEG activity into commands, even without a voluntary intention from the users. Hence, a supervisor is always required to set the application up. This mode is usually unpractical, and should be avoided if the system is intended to have possibilities to work in a real situation (Pinegger et al., 2015). Self-paced BCIs cannot be achieved without a proper *asynchrony* management, which makes systems always available and gives users the power to control when the BCI output should be produced (Wolpaw Jonathan R., 2012).

1.5 Motor disabilities

Validation of BCIs with disabled people is crucial for identifying needs, testing and determining the reliability of the systems. Among the wide variety of motor disabilities, people with locked-in syndrome (LIS) were identified as the most immediate target population for BCIs. In this condition, the patient is conscious but cannot communicate or move due to a complete paralysis of all limbs and the last cranial nerves (Hochberg Leigh R., 2012). Note that LIS is not a disease, but a state resulted from different clinical etiologies. However, it is not necessary to reach that point, but also people with any kind of impaired mobility may be

target users for BCIs. In general, motor disabilities of most BCI users are often caused by neurodegenerative diseases, traumas, muscle disorders, or any illness that impairs the neural pathways that control muscles or the muscles themselves (Wolpaw et al., 2002).

All the assistive BCI applications that are presented in this Doctoral Thesis have been evaluated with target users, in order to provide an accurate measurement of performance in a real environment. All subjects presented different motor disabilities caused by one of the following diseases:

1) **Multiple sclerosis (MS).** MS is the most common autoimmune disease that affects the CNS. In MS, the immune system damages the myelin of brain and spinal cord nerve cells, causing an inflammation that disrupts the ability of neurons to communicate. This results in a variety of symptoms, including motor problems, cognitive deficits or psychiatric disorders. Indeed, up to 60% of the patients develop a motor disability before twenty years after the onset (Compston and Coles, 2008). In rare cases the disease may be terminal, although most people with MS have a normal life expectancy (Compston and Coles, 2008).

2) **Stroke.** Also known as cerebrovascular accident, the stroke is the major cause of disability and second most common cause of death worldwide. In this condition, brain cells die because of insufficient blood flow, caused either by an ischemic (i.e., lack of blood flow) or a hemorrhagic (i.e., bleeding) origin. Both lead to functional brain damage in the affected regions, which can be permanent (Donnan et al., 2008).

3) **Spinal cord injury (SCI).** An SCI is a damage to the spinal cord that results in temporary or permanent loss of mobility or feeling. The injury (caused by trauma or disease) can be either *complete* if a total loss of function is presented, or *incomplete* if some nerves are able to pass the damaged area. Depending on the degree of affectation, the location, and the rehabilitation procedure, long-term SCI outcomes range from complete recovery to permanent paraplegia or quadriplegia (McDonald and Sadowsky, 2002).

4) **Cerebral palsy (CP).** CP is the most common movement disorder in children. It reflects an abnormal brain development that usually leads to spasticity (i.e., stiff muscles), dyskinesia (i.e., uncontrollable movements) and ataxia (i.e., poor coordination) (NINDS, 2013). Often, the cause is unknown; though is closely related to problems that occur during pregnancy,

childbirth or immediately after it. There is no cure for CP and thus, people with CP require long-term care, including medications, therapies and even surgical procedures (NINDS, 2013).

5) **Muscular dystrophies (MD).** MD comprises a group of muscle diseases that lead to progressive weakness and loss of muscle mass. Although there are several types of MD, all of them are considered inherited genetic disorders; i.e., caused by mutations on genes involved in making muscle proteins. There is no cure for MD and most people eventually become unable to walk. However, symptoms may be monitored and relieved by medication, physical therapy and surgery (Mercuri and Muntoni, 2013).

6) **Friedreich's ataxia (FRDA).** FRDA is an inherited neurodegenerative disease in which the production of the protein fraxatin is hindered, resulting in cellular damage and death. The ataxia is eventually caused by the degeneration of the spinal cord nerves. Symptoms, which usually starts in childhood, progressively worsen until people lose the ability to stand or walk. FRDA has no cure; thus, patients rely on physical and occupational therapies the rest of their lives (Delatycki et al., 2000).

1.6 State of the art

The main contributions of this Doctoral Thesis are related with the asynchrony of the BCIs, the channel selection problem and the development of assistive applications (i.e., web browser, social networking app). Therefore, the next subsections reflect a comprehensive state of the art revision of these topics:

1.6.1 Asynchrony

As mentioned in section 1.4.4, synchrony is one of the main limitations of current P300-based BCIs. Conventional synchronous BCIs are typically associated with visual *oddball paradigms*, in which random flashings elicit P300 potentials and allows the system to discriminate among different commands. As the BCI cannot monitor users' attention, it will continue selecting random commands even if the user is ignoring the stimuli. Thus, recent investigation have been focused on providing a real asynchronous control of the system; i.e., to discriminate between control (attending) and non-control (ignoring) states.

Based on the hypothesis that P300 potentials are more easily discriminated when users are attending to the stimuli, many studies rely on a threshold derived from classifier's scores, which are expected to be higher in the control state. These scores were obtained from support vector machines (SVM) (He et al., 2017; Li et al., 2013; Zhang et al., 2008) or linear discriminant analysis (LDA) (Alcaide-Aguirre et al., 2017; Aloise et al., 2011; Aref and Huggins, 2012; Breitwieser et al., 2016; Martínez-Cagigal et al., 2017, 2019a; Tang et al., 2018) classifiers. Aydin et al. (2018) proposed the use of labels instead of scores for designing different criteria to define the non-control state. Some studies proposed modifications of the paradigm to elicit SSVEPs; using P300 potentials to determine the target command, and SSVEPs to identify the non-/control states (Li et al., 2013; Panicker et al., 2011). Similarly, other studies took a step further, hypothesizing that the stimulation already produces residual SSVEPs in control states, provided flashings occur at a fixed rate. Hence, spectral methods were also proposed as features, such as relative powers (Ma and Qiu, 2018), sums of spectral components (Pinegger et al., 2015) or canonical correlation analysis (CCA) (Santamaría-Vázquez et al., 2019). Further modifications for hybrid BCIs have been also proposed, involving MI tasks via SMR (Yu et al., 2017) or tactile ERPs (Breitwieser et al., 2016).

1.6.2 Channel selection

Channel selection optimization is usually overlooked in BCI literature due to its inherent complexity. Nevertheless, a channel selection procedure is beneficial to reduce the dimensionality of the data, the cost of the system, the power consumption and to increase users' comfort. According to the aforementioned limitations of current BCIs (see section 1.4.4), this problem involves the hardware and reliability aspects of the systems.

Generally, P300-based studies use the 8-channel combination proposed by Krusienski et al. (2008), mainly located over parieto-occipital regions. This approach is valid as a general rule of thumb, but ignores the intrinsic inter-subject variability of ERP responses. For this reason, many recent studies were aimed at testing novel metrics to find the most appropriate channels for each subject. Note that an exhaustive search is intractable, due to the enormous amount of possible combinations (i.e., 2^{N_c} for a N_c-channel cap). Meta-heuristics have demonstrated excellent performances solving complex optimization problems, especially swarm intelligence and evolutionary computation approaches. In particular, particle swarm optimization (PSO), a single-objective algorithm, has been successfully

applied in P300-based BCIs (Arican and Polat, 2020; Gonzalez et al., 2013, 2014; Jin et al., 2010; Perseh and Sharafat, 2012). However, a BCI-oriented channel selection should optimize two objectives simultaneously: (1) to maximize the performance, and (2) to minimize the number of channels. As a preliminary work, we used a weighted aggregation approach to merge both objectives into a single expression and adapted 5 single-objective approaches to the discrete domain, including PSO, bees algorithm (BA), artificial bee colony (ABC), binary ant system (BAS) and firefly algorithm (FA) (Martínez-Cagigal and Hornero, 2017a,b). Nevertheless, simultaneous optimization was not explored; indeed, the application of multi-objective algorithms in P300-based BCI channel selection is more limited. To the best of our knowledge, only Kee et al. (2015) applied a non-sorting genetic algorithm II (NSGA-II), while Chaurasiya et al. (2017) used a multi-objective binary differential-evolution algorithm (MOBDE), achieving promising results. In spite of their scarce application, evolutionary computation applied is a growing research field that integrates many different algorithms that have not been fully exploited yet in this context. Moreover, their application in P300-based BCIs is not straightforward, due to the dichotomy of the problem (e.g., select or reject each channel). The discretization of these meta-heuristics must be performed carefully, by means of adapting inner equations and employing binary operators (e.g., crossover, mutation, transfer functions, etc.).

1.6.3 BCI web browsers

As aforementioned, the most straightforward application of BCI systems is to improve the quality of life of the severely disabled, for instance, by replacing the control of some devices to make them assistive and favor their accessibility. Owing to the advance of Internet in the last decades, it is natural to think whether a BCI web browser could be feasible. Nowadays, web browsers are designed to be controlled by a keyboard and a mouse, but not with a small amount of control signals.

In this regard, several previous attempts aimed at controlling a BCI web browser were reported in the literature. The first ones employed SCP or SMR as control mechanisms, using dichotomous decision trees to select or reject commands (Bensch et al., 2007; Karim et al., 2006). Notwithstanding their usefulness as the precursors of BCI-based web browsers, the binary selection strategy was slow and required a supervisor to adjust several parameters (e.g., reading pauses, address book entries, etc.). Furthermore, users needed to spend a lot of time to

learn to control SCP and SMR signals, and it was not guaranteed that they could eventually control them with enough ability to experiment a real command of the system (Bensch et al., 2007; Karim et al., 2006). Mugler et al. (2010) overcome the selection slowness by developing a P300-based system, in which commands were selected using the *oddball* paradigm. Page links were tagged with an alphanumeric coding, allowing its selection provided the code was entered using P300 selections. Sirvent Blasco et al. (2012) also used a P300-based approach, but the selection of links was made by simulating the movement of the mouse trough commands that shifted its position a fixed amount of pixels in different directions. A hybrid approach was also developed by Yu et al. (2012), in which the horizontal and vertical movements of the mouse were controlled by SMR and P300 potentials, respectively. Nonetheless, these approaches were synchronous and thus, it was required to include a "read mode" command to pause the stimulation when the user wanted to calmly read a web page . This pause, however, was fixed for a predefined number of seconds, which could result too long or too short a time for the user (Mugler et al., 2010; Sirvent Blasco et al., 2012). For a truly free surfing, this synchronous mode is impractical and cannot be applied in a real setup. To the best of our knowledge, there are not asynchronous-based BCI web browsers in the literature. Noteworthy, Karim et al. (2006), Bensch et al. (2007) and Mugler et al. (2010) tested their proposals with 1, 4 and 3 ALS patients, respectively; reaching averaged accuracies of 80% and 72% (results of Bensch et al. (2007) were not reported). Accuracies of healthy subjects, by contrast, reached a mean of 93% (Mugler et al., 2010; Sirvent Blasco et al., 2012; Yu et al., 2012), which compromises the generalization of the results and the feasibility of those studies that lacked of validation with motor-disabled patients.

1.6.4 Mobile BCIs

Other BCI application could be focused on controlling some functionalities of the smartphones. Currently, these devices have more than 4.9 billion of unique users in the world, becoming an essential aspect of our daily lives (Kemp, 2018). Although their functionalities cover managing finances, reading news, watching videos, shopping, playing games or searching for information, among other uses; more than the 56% of the time spent is dedicated to socializing via social media and instant messaging (Ipsos MORI and Google, 2017). In spite of this development, smartphones are still restricted for motor-disabled people than cannot control their fingers or hands accurately. Therefore, whether a BCI controlled smartphone could

bridge the accessibility gap in these trending devices is a reasonable concern.

Despite the popularity of smartphones or tablets nowadays, studies that aim at controlling their functionalities using BCIs are scarce in the literature. These studies, tested with healthy subjects (HS), are limited to dial numbers in cell phones (Chi et al., 2012; Wang et al., 2011), call contacts (Campbell et al., 2010; Wang et al., 2011), accept incoming calls (Katona et al., 2014), play a simple racing game (Wu et al., 2014), spell words (Campbell et al., 2010; Elsawy et al., 2017; Obeidat et al., 2017), or open pre-installed apps and visualize the image gallery (Elsawy and Eldawlatly, 2015). Among them, the works of Wang et al. (2011) and Chi et al. (2012) used SSVEPs as control signals, reaching averaged accuracies of 95.90% (10 HS) and 89.00% (2 HS), respectively. Others used a commercial control signal from the Neurosky cap, achieving an accuracy of 75.00% (5 HS) (Katona et al., 2014; Wu et al., 2014). The rest of them used P300 potentials, reaching accuracies of 88.89% (3 HS) (Campbell et al., 2010), 83.34% (6 HS) (Elsawy and Eldawlatly, 2015), 64.17% (6 HS) (Elsawy et al., 2017) and 90.00% (14 HS) (Obeidat et al., 2017). Nevertheless, to the best of our knowledge, none of those studies has been focused on making social apps accessible, nor providing a high-level control of a smartphone or tablet. In addition, results show a poor performance of HS compared with those reported in other P300-based BCI studies, likely due to simple EEG acquisition equipment and signal processing pipelines, leaving room for improvement. Furthermore, they have not tested their proposals with motor-disabled people, which are presumably the target users of their applications. From these state-of-the-art revision it is thus clear that a BCI system to control socializing functionalities of a smartphone could be a novel contribution to the literature, as well as a suitable assistive application to meet the daily communication needs of the motor-disabled.

Once the main topics of this Doctoral Thesis have been introduced, the rest of the document is organized as follows. Chapter 2 enumerates the hypotheses that have motivated each study, including the overall purpose and the specific objectives. Databases and EEG acquisition procedure are detailed in the chapter 3. Afterward, chapter 4 describes the methodology, including signal pre-processing, feature extraction, selection and classification, channel selection, performance assessment and statistical analysis. Main results are shown in the chapter 5, which are further discussed in the chapter 6, enumerating the current limitations. Finally, the contributions of this Doctoral Thesis, as well as the final conclusions, are detailed in the chapter 7. The last sections are intended to complement this doc-

ument by including: the papers of the compendium of publications (appendix A), the scientific achievements achieved during the Ph.D. (appendix B), and a brief summary in Spanish (appendix C).

Chapter 2

Hypothesis and objectives

As it has been previously shown, developing reliable BCI systems for improving the quality of life of the severely disabled has become a major concern in the last decades. Therefore, the proposals developed in this Doctoral Thesis are focused on providing asynchronous and channel selection signal processing methodologies and assistive applications, intended to provide a real use of P300-based BCIs by motor-disabled people. Thereby, the hypotheses that have motivated each study, as well as the joint main hypothesis that justify the present Doctoral Thesis are declared in section 2.1. On the other hand, section 2.2 introduces the main objective and the specific objectives that must be fulfilled to reach it.

2.1 Hypothesis

Despite the growing interest of scientific literature in BCI systems in the last years, the current limitations of these systems have relegated their application to laboratories. Consequently, studies usually overlook the potential utilization of the BCI on a real situation and focus on merely academic purposes. Thus, a naive hypothesis can be formulated as a starting point: *limitations of current BCIs that hinder their application outside the laboratories can be lessened*. Notwithstanding its usefulness as an introductory step, this high-level statement does not suffice to approach any particular investigation by its own. Therefore, additional lower level hypotheses need to be examined for reaching that point.

As stated in subsection 1.4.4, current BCIs should mainly face the following limitations: (1) reliability, (2) synchrony, (3) hardware, and (4) validation. In general, (1) reliability may be addressed by using control signals with a proper

inter-subject generalization ability and a reduced calibration. In other words, signals that do not depend on the ability of users to self-regulate their own activity, such as P300 evoked potentials (as discussed in subsection 1.4.3). Addressing the (2) synchrony is absolutely essential to provide a comprehensive self-paced control of the BCI. Asynchrony can only be achieved if the BCI is able to determine whether the user wants to deliver a command (i.e., control state) or not (i.e., non-control state). Thereby, it has been assumed that *novel feature extraction and classification approaches are able to reach an asynchronous control.* Since it can be thought that control state is more demanding than non-control, it is particularly hypothesized that *entropy-based measurements can characterize the regularity of non-/control EEG signals and provide complementary information for discriminating between both states.* (3) Hardware limitations may be mainly addressed through electronic improvements (e.g., dry electrodes, wireless communication, battery autonomy, etc.), although some software approaches can contribute too. For instance, reducing the number of electrodes would reduce the cost of the system, power consumption and increase users' comfort. Due to their great abilities to solve complex optimization problems, it has been hypothesized that *meta-heuristics based on evolutionary computation are able to optimize the channel set for each user in P300-based BCIs.* Similarly, it has been assessed whether *these approaches effectively avoid the curse of dimensionality and maximize the P300 potential detection accuracy*, somehow contributing to the (1) reliability issue as well. Regarding the last limitation, every single BCI intended to be used by the severely disabled should be (4) validated with a target population for guaranteeing a proper performance, and not only with healthy subjects. Nowadays, there are still numerous applications and devices whose accessibility is greatly limited for motor-disabled people (e.g., surfing the Internet, accessing social networks). It has been thus hypothesized whether *a P300-based asynchronous BCI web browser is able to provide a reliable Internet surfing for motor-disabled people.* By extension, it has been also assessed if *a P300-based asynchronous BCI can provide a comprehensive control of smartphone-based social networks for the motor-disabled*; including Twitter –a microblogging service–, and Telegram –a cloud-based instant messaging application–.

These statements are the main hypotheses that form the core of the present Doctoral Thesis, which can be merged into the following global hypothesis:

> "*P300-based BCI systems may be oriented toward a real use outside laboratories by applying novel signal processing methodologies (e.g., asynchronous management, channel selection) and developing asyn-*

chronous applications (e.g., web browser, social networks)."

2.2 Objectives

The general goal of this Doctoral Thesis is to design, develop and evaluate novel signal processing methodologies and assistive applications toward a real use of P300-based BCIs by motor-disabled people. These methodologies were focused on optimizing channel sets and reaching a fully self-paced control of the systems; whereas the assistive applications were aimed at bridging the accessibility gap in web browsers and smartphone-based social networks. In order to achieve the main objective, the following specific objectives arise:

I. To review the bibliography and state-of-the-art related to non-invasive BCI systems, putting special emphasis in channel selection and asynchronous techniques, as well in development of assistive applications.

II. To build a database of non-/control state EEG recordings for approaching the asynchronous problem in P300-based BCIs and to recruit a large population of motor-disabled users (including MS, stroke, SCI, CP, MD, and FRDA), as well as their socio-demographic data, to validate the assistive applications.

III. To implement the most appropriate methods to optimize channel sets, discriminate between asynchronous states, and identify P300 potentials; and to investigate the suitability of further improvements.

IV. To design and develop two P300-based assistive asynchronous systems for controlling: (1) a web browser, and (2) smartphone-based social networks.

V. To evaluate the ability of the selected methods for optimizing the channel set of each user and reaching a self-paced control of P300-BCIs by testing them in both public and previously recorded databases, as well as to validate the developed P300-based assistive asynchronous BCIs with the recruited motor-disabled population, as well as with healthy subjects.

VI. To conduct statistical analysis of results for evaluating the suitability of each proposal, and for characterizing performances depending on the population; and to compare and discuss the results to draw appropriate conclusions, including a comprehensive comparison with previous state-of-the-art studies.

VII. To disseminate the main results and conclusions in JCR indexed journals, as well as in book chapters and international and national conferences.

Chapter 3

Subjects and signals

In this chapter, different databases and acquisition setups that play an important role in the compendium of publications are revised. Firstly, section 3.1 details demographic and clinical characteristics of the subjects that compose the databases of each study. The three different oddball paradigms employed to elicit P300 potentials are presented in section 3.2. Finally, section 3.3 describes the acquisition setups, including EEG equipment, electrode locations and sampling rate.

3.1 Subjects

As the purpose and validation procedure of every conducted study vary, different databases and subject pools were used in this Doctoral Thesis. A general specification of these databases is shown in the table 3.1. Details of each of them, such as number of subjects, sex, age and degree of disability of the participants are indicated next:

1) **Entropy-based asynchrony study.** As the study was aimed at characterizing both non-/control states, an asynchronous database was recorded. This database was composed by 10 HS (6 males, 4 females, mean age: 25.70 $\pm$ 3.09 years) that were asked for ignoring and attending the row-col paradigm (RCP), using a 16-channel EEG cap (Martínez-Cagigal et al., 2019b).

2) **Channel selection study.** For the analysis of meta-heuristics to optimize the number of channels, three different public databases were used (Martínez-Cagigal et al., 2020). All of them were recorded from HS with high-density EEG caps, but employed different stimulation paradigms. The first one

Table 3.1: Database specifications for each study.

Study	Disabled patients	Healthy subjects	Paradigm	No. channels
Martínez-Cagigal et al. (2019b)	0	10	RCP	16
Martínez-Cagigal et al. (2020)	0	2	RCP	64
	0	13	CS	63
	0	12	RSVP	61
Martínez-Cagigal et al. (2017)	16 MS	5	RCP	8
Martínez-Cagigal et al. (2019a)	18*	10	RCP	8

MS: multiple sclerosis, RCP: row-col paradigm, CS: center speller, RSVP: rapid serial visual presentation. * 1 stroke, 2 spinal cord injury, 5 Friedreich's ataxia, 5 cerebral palsy, 2 muscular dystrophy.

is the 'BCI Competition III: dataset II', recorded from 2 HS using a 64-channel EEG cap with the RCP (Blankertz et al., 2006). The second one is the 'Center Speller (008-2015)' dataset, recorded from 13 HS (8 males, 5 females, mean age: 27 years) using a 63-channel EEG cap with the center speller paradigm (Treder et al., 2011). The third one is the 'RSVP Speller (010-2015)' dataset, recorded from 12 HS (6 males, 6 females, mean age: 29.17 years) using a 61-channel cap with the rapid serial visual presentation paradigm (Acqualagna and Blankertz, 2013).

3) **Web browser study.** A reliable validation of an asynchronous BCI web browser for the severely disabled requires recruiting target users, in addition to the HS pool. Therefore, the database was composed by 16 MS patients (10 males, 6 females, mean age: 42.06 ± 7.47 years) and 5 HS (all males, mean age: 26.00 ± 4.58 years) using a 8-channel cap, whose demographic and clinical data are detailed in the table 3.2 (Martínez-Cagigal et al., 2017).

4) **Smartphone-based social networks study.** Similarly, the subject pool that evaluated the smartphone-based social networks study was composed by 18 motor-disabled subjects (1 stroke, 2 SCI, 5 FRDA, 5 CP, 2 MD; 11 males, 7 females, mean age: 47.28 ± 9.68 years) and 10 HS (8 males, 2 females, mean age: 26.10 ± 3.45), using a 8-channel cap. Their demographic and clinical data are depicted in the table 3.3 (Martínez-Cagigal et al., 2019a).

All subjects gave their informed written consent for participating in the studies, whose protocols were approved by the local ethics committee and conformed to the declaration of Helsinki. Motor-disabled participants of Martínez-Cagigal et al. (2017) and Martínez-Cagigal et al. (2019a) were recruited by the National Refer-

Table 3.2: Participants' demographic and clinical data for the web browser study.

	User	Sex	Age	DMD	DCA	DSA
Healthy	CH01	M	23	-	-	-
	CH02	M	31	-	-	-
	CH03	M	23	-	-	-
	CH04	M	31	-	-	-
	CH05	M	22	-	-	-
MS Patients	CP01	F	30	Non-existent	Very high	Very high
	CP02	M	31	Non-existent	High	Very high
	CP03	M	43	Mild	Very high	High
	CP04	F	47	Moderate	Normal	High
	CP05	M	56	Moderate	Low	Very low
	CP06	F	32	Non-existent	Normal	Normal
	CP07	M	35	Non-existent	Very high	Very high
	CP08	M	41	Non-existent	High	High
	CP09	F	49	Non-existent	Normal	Very high
	CP10	M	44	Mild	Normal	Low
	CP11	F	41	Moderate	Normal	High
	CP12	M	43	Moderate	Very high	Normal
	CP13	M	44	Non-existent	High	High
	CP14	M	52	Moderate	Very high	Normal
	CP15	F	38	Non-existent	Normal	High
	CP16	M	47	Moderate	Normal	Normal

DMD: degree of motor disability, DCA: degree of cognitive ability, DSA: degree of sustained attention, M: male, F: female (Martínez-Cagigal et al., 2017).

ence Centre on Disability and Dependence, located in San Andrés del Rabanedo (León, Spain).

3.2 Oddball paradigms

As discussed in section 1.4.3, P300 evoked potentials are endogenous components elicited by rare stimuli that transgress the user's expectations. The experimental procedure that is required to produce these potentials is known as *oddball paradigm*, which must keep to three main attributes: (1) there are two different classes of stimuli; (2) stimuli that fall into one class are much more frequent than for the other class; and (3) the task compel the subject to classify each stimulus into one of the two classes. In fact, the absence of a stimulus may be a class if the setup satisfies the aforementioned conditions (Sellers Eric W., 2012). In the BCI field, visual *oddball paradigms* are widely used, leading to the proposal of many variants. Regarding this Doctoral Thesis, three different paradigms were used:

1) **Row-Col Paradigm (RCP).** Probably the most common application of the *oddball paradigm* in BCI systems (Farwell and Donchin, 1988). As shown

Table 3.3: Participants' demographic and clinical data for the smartphone-based social networks study.

	User	Sex	Age	DMD (%)	Disease
Healthy	DH01	M	25	-	-
	DH02	M	25	-	-
	DH03	M	24	-	-
	DH04	M	25	-	-
	DH05	M	25	-	-
	DH06	M	32	-	-
	DH07	M	24	-	-
	DH08	M	25	-	-
	DH09	F	23	-	-
	DH10	F	33	-	-
Motor-disabled subjects	DP01	F	48	90%	Stroke
	DP02	M	46	80%	Spinal cord injury
	DP03	F	38	93%	Friedreich's ataxia
	DP04	M	39	98%	Spinal cord injury
	DP05	F	49	78%	Friedreich's ataxia
	DP06	M	31	76%	Cerebral palsy
	DP07	M	52	99%	Cerebral palsy
	DP08	M	44	90%	Friedreich's ataxia
	DP09	M	47	69%	Cerebral palsy
	DP10	M	67	87%	Cerebral palsy
	DP11	M	62	86%	Muscular dystrophy
	DP12	M	47	90%	Muscular dystrophy
	DP13	F	66	94%	Friedreich's ataxia
	DP14	F	40	88%	Friedreich's ataxia
	DP15	M	38	98%	Spinal cord injury
	DP16	M	50	80%	Spinal cord injury
	DP17	F	42	89%	Cerebral palsy
	DP18	F	45	84%	Spinal cord injury

DMD: degree of motor disability, M: male, F: female (Martínez-Cagigal et al., 2019a).

in figure 3.1(A), a matrix containing alphanumeric characters or commands is displayed. Rows and columns of the matrix flash in a randomized fashion until every single row and column has been illuminated, which is known as a sequence. The user, who must focus on the desired command, will elicit a P300 potential whenever the column or the row that contain the target command are flashed.

2) **Center Speller (CS).** The CS paradigm was proposed as a gaze-independent variation of RCP for patients with impaired eye movements (Treder et al., 2011). As shown in figure 3.1(B), CS exploits feature attention based on colors and shapes that are sequentially presented on the center of the screen. Commands are grouped into six different shapes that

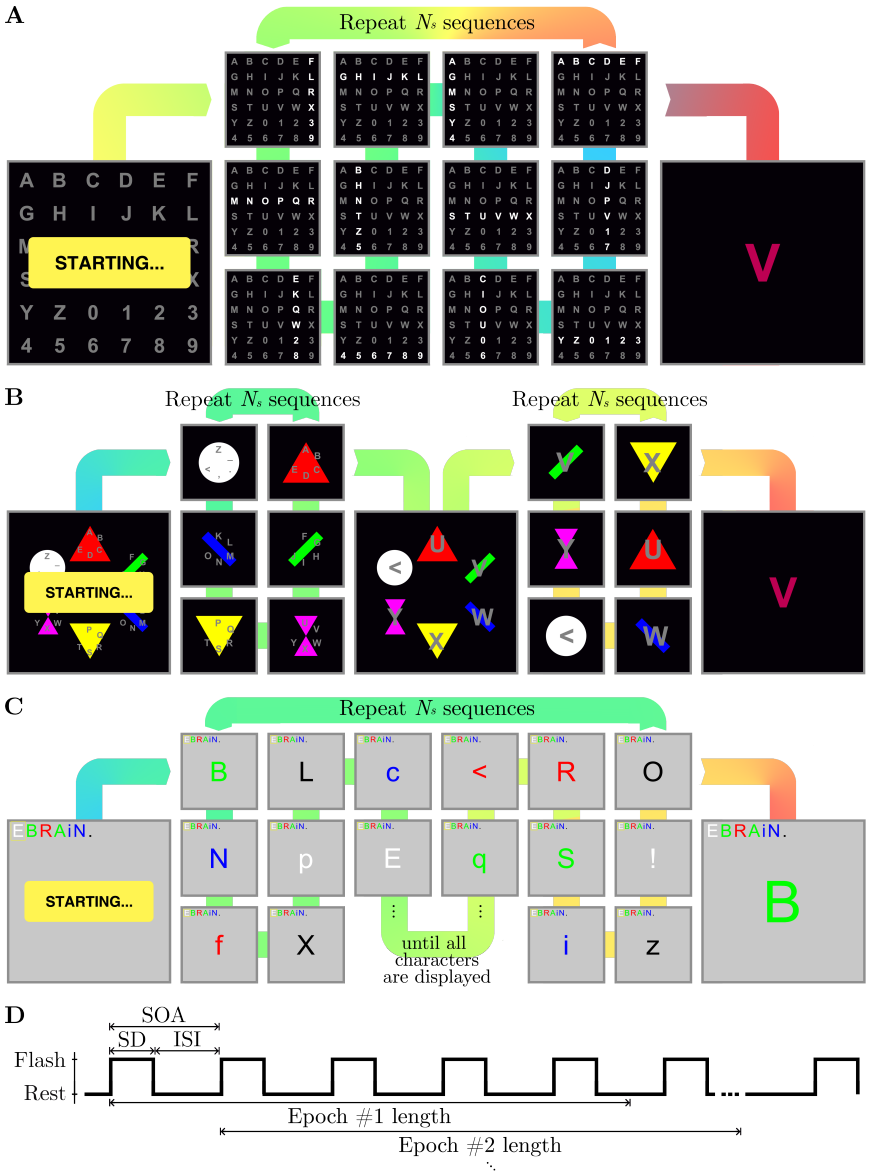


Figure 3.1: Visual *oddball paradigms* used in this Doctoral Thesis. (A) Series of flashings in the RCP paradigm. Since the matrix dimensions are 6×6, a total of 12 randomized flashings are required to highlight each row and column. The procedure is repeated N_s times to increase the number of observations and strengthen the P300 wave. (B) Series of flashings in the CS paradigm. Output commands are grouped in 6 sets with different shapes and thus, the selected command is determined through two different selection stages. Note that the N_s repetitions are performed in each stage separately. (C) Series of flashings in the RSVP paradigm. Characters are sequentially displayed on the center of the screen, using color and uppercase/lowercase variations. (D) Time course of a sequence of flashings in visual *oddball paradigms* such as RCP, CS or RSVP.

Table 3.4: Oddball parameters for each database.

Study	Paradigm	$N_s^{\dagger}$	Timings		
			SD	**ISI**	**SOA**
Martínez-Cagigal et al. (2019b)	RCP	15	75 ms	100 ms	175 ms
Martínez-Cagigal et al. (2020)	RCP	15	100 ms	75 ms	175 ms
	CS	10	100 ms	100 ms	200 ms
	RSVP	10	N.r.	N.r.	83 ms
Martínez-Cagigal et al. (2017)	RCP	15	62.5 ms	$\mathcal{U}$(125, 250) ms	$\mathcal{U}$(187.5, 312.5) ms
Martínez-Cagigal et al. (2019a)	RCP	15	62.5 ms	$\mathcal{U}$(125, 250) ms	$\mathcal{U}$(187.5, 312.5) ms

SD: stimulus duration, ISI: inter-stimuli interval, RCP: row-col paradigm, CS: center speller, RSVP: rapid serial visual presentation, N.r.: not reported (Treder et al., 2011), $\mathcal{U}$ randomized within uniform distribution.
† Number of sequences of the training data.

flash according to the *oddball paradigm*, allowing the user to select them in two stages: (i) the group, (ii) the specific command. Unlike the RCP, this paradigm does not require an accurate control of the gaze, since flashings are presented on the center of the screen.

3) **Rapid Serial Visual Presentation (RSVP).** The RSVP paradigm was developed to exploit the foveal visual field and avoid eye movements (Acqualagna and Blankertz, 2013). As shown in figure 3.1(C), RSVP depicts symbols in the center of the screen in a serial manner. In order to favor the identification of the symbols, half of the characters are uppercase and the other half lowercase, using 5 different colors. A sequence is thus finished when all symbols have been presented.

The procedure is repeated N_s times to get enough observations for assuring the reliability of the P300 potential. Note that few flashings will elicit a P300 (i.e., infrequent class) in comparison with those that are not associated to the target command (i.e., frequent class), leading to ratios of 1:6 in both examples. Generally, calibration data is recorded with $10 \leq N_s \leq 15$ sequences; being reduced to an understandable number after classifier training. Of course, flashings should be randomized inside a sequence; otherwise, the user could expect their onsets, violating the core principle of P300 elicitation: an unexpected stimulation.

Figure 3.1(D) depicts the time course of a sequence of flashings. The stimulus onset asynchrony (SOA), i.e. the elapsed time between two consecutive onsets; is composed by the stimulus duration (SD) and the inter-stimuli interval (ISI). Although there is not a general consensus about the fixation of these parameters,

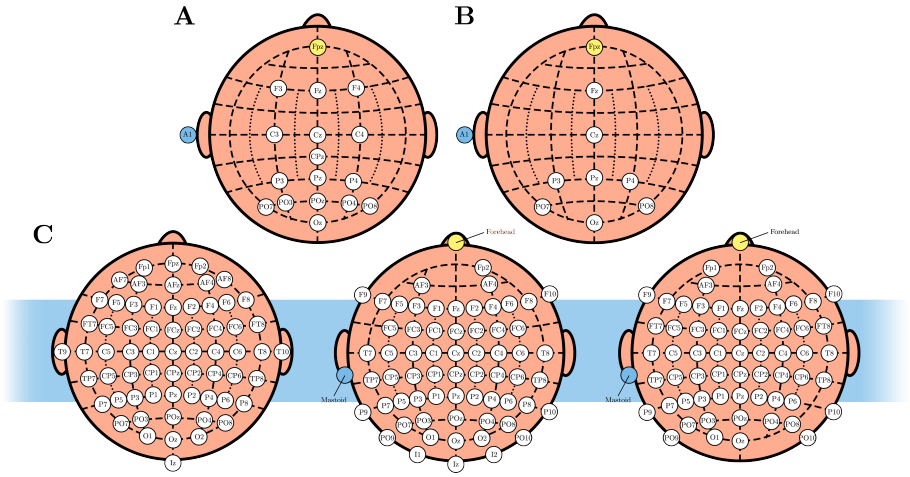


Figure 3.2: Electrode montages used in this Doctoral Thesis, according to the International System 10–10: (A) asynchrony study (Martínez-Cagigal et al., 2019b); (B) application studies (Martínez-Cagigal et al., 2017, 2019a); (C) RCP (left), CS (center), and RSVP (right) databases of the channel selection study (Martínez-Cagigal et al., 2020). Ground and reference electrodes are marked in yellow and blue, respectively.

SOA usually takes values in the range of 175–200 ms (Acqualagna and Blankertz, 2013; Blankertz et al., 2006; Treder et al., 2011; Wolpaw and Wolpaw, 2012a). Occasionally, ISI varies among sequences to hinder users' anticipation. Table 3.4 indicates the values that have been used in this Doctoral Thesis. It is noteworthy that the epoch length used to classify both classes typically contains several flashings and thus, the output ERP is a superposition of delayed P300 waves.

3.3 Acquisition setup

As the databases vary, the acquisition setup also varied among the studies. Figure 3.2 shows the electrode montages for each study, according to the International System 10–10 (see section 1.4.2). EEG signals for the asynchrony (Martínez-Cagigal et al., 2019b), web browser (Martínez-Cagigal et al., 2017) and social networks (Martínez-Cagigal et al., 2019a) studies were recorded with a g.USBAmp (*g.tec Medical Engineering*, Austria), at a sampling rate of 256 Hz and using 16, 8 and 8 active electrodes, respectively. Data was referenced to the earlobe, using the channel Fpz as a ground. Conversely, the public databases that were used in the channel selection study (Martínez-Cagigal et al., 2020) differ: RCP database was recorded using 64 channels at 240 Hz (equipment, reference and ground not

reported) (Blankertz et al., 2006); CS and RSVP databases were recorded via 63 and 61 active electrodes connected to BrainAmp systems (*Brain Products*, Germany) at 1000 Hz, referenced to the left mastoid and placing a ground electrode over the forehead (Acqualagna and Blankertz, 2013; Treder et al., 2011).

Chapter 4

Methods

This chapter describes the methods that have been applied through the compendium of publications. Signal conditioning and channel selection algorithms are detailed in sections 4.1 and 4.2, respectively; followed by the standard EEG signal processing pipeline: feature extraction (section 4.3), selection (section 4.4) and classification (section 4.5). Afterward, section 4.6 is devoted to the asynchrony management. Different methods to assess the performance of the proposed algorithms are detailed in section 4.7. Finally, section 4.8 provides a brief description of applied statistical analyses.

4.1 Signal pre-processing

After the EEG acquisition, a signal conditioning or pre-processing step must be performed to remove artifacts and enhancing temporal, spectral or spatial characteristics of the signals:

1) **Frequency filtering.** As EEG frequencies above 40 Hz have a very low signal-to-noise ratio (see section 1.4.2) and amplifier's low-frequency drifts are sometimes present, a band-pass filtering is essential for BCI applications (Krusienski Dean J., 2012). In this case, signals were band-pass filtered between 0.1–60 Hz and then, a 50 Hz notch filter was applied to remove the power interference. For the offline studies, finite impulse response (FIR) filters were used (Martínez-Cagigal et al., 2019b, 2020); for the online ones, infinite impulse response (IIR) filters were applied (Martínez-Cagigal et al., 2017, 2019a).

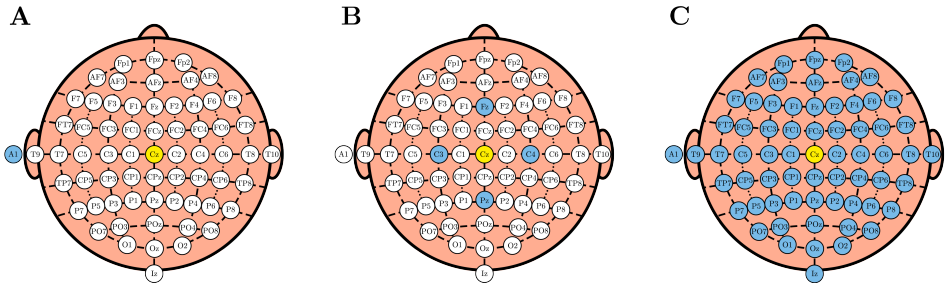


Figure 4.1: Spatial filtering techniques in EEG-based BCIs. The mean of blue electrodes would be subtracted from the current channel (e.g., Cz), depicted in yellow. (A) Ear reference. (B) Laplacian filtering, (C) CAR filtering.

2) **Spatial filtering.** Spatial filters are fundamentally used to improve source localization, to enhance particular channels or to remove certain artifacts. Considering the latter, a common average reference (CAR) filtering was applied by subtracting, for each channel, the mean value of all of them, as shown in figure 4.1(C). Hence, $\boldsymbol{x}_c = \boldsymbol{x}_c - \frac{1}{N_c}\sum_{i=1}^{N_c}\boldsymbol{x}_i$, where $\boldsymbol{x}_c$ is the signal for the electrode c, and N_c is the number of channels (Krusienski Dean J., 2012). In this way, artifacts that are common to all electrodes, such as power interference, are minimized. CAR provides a biased estimate of reference-independent potentials as well, whose approximation error decreases in high-density electrode montages.

4.2 Channel selection

The stimuli repetitions that are required to elicit a reliable P300 response (see section 3.2) constitute high-dimensional data liable to produce an over-fitting of the classifier. This phenomenon, known as *the curse of dimensionality*, may produce spoiled performances if data are not properly processed by means of feature extraction and/or selection methods, regularized classifiers or channel selection procedures. The latter is often overlooked in BCI literature, due to its inherent complexity; and thus it is common to see a 8-channel set in P300-based BCIs as a general rule of thumb (Krusienski et al., 2008). However, among these approaches, channel selection procedures are the only that are able to reduce the cost of the system, the power consumption and to increase users' comfort. For this reason, we studied the ability of several single- and multi-objective optimization approaches for selecting the most appropriate channel set for each user Martínez-Cagigal et al.

(2020). We also proposed a novel meta-heuristic that overcame their limitations, while the rest of the studies of this Doctoral Thesis used the Krusienski's set (Martínez-Cagigal et al., 2017, 2019a), or an extended version of it (Martínez-Cagigal et al., 2019b).

Dealing with the problem of selecting the most suitable set of channels is not trivial. For an N_c-channel cap, there are 2^{N_c} possible subset combinations of channels to consider, making an exhaustive search intractable in practice. Channel selection is therefore viewed as a complex optimization problem, where iterative meta-heuristics are applied to find an optimal (or suboptimal but suitable) solution in a reasonable amount of time (Bozorg-Haddad et al., 2017). Here, a solution is represented as $\boldsymbol{s} = [s_1, s_2, \ldots, s_{N_c}], s_i \in \{0, 1\}$, where 1 and 0 values indicate the selection or rejection of a specific channel i, respectively. Moreover, there must be a way to assess the *fitness* of each solution for achieving an objective function. In a P300-BCI, two main objectives should be pursued: (1) to maximize the performance of the system, and (2) to minimize the number of EEG channels; which may be modeled as:

$$\min F(\boldsymbol{s}) = \begin{cases} f_1(\boldsymbol{s}) = 1 - \overline{AUC}(\boldsymbol{s}) \\ f_2(\boldsymbol{s}) = \sum_{i=1}^{N_c} s_i \end{cases}, \tag{4.1}$$

where $f_1(\boldsymbol{s})$ and $f_2(\boldsymbol{s})$ are the first and second objective, respectively. $\overline{AUC}(\boldsymbol{s})$ denotes the area under receiving-operating characteristic (ROC) curve for the solution $\boldsymbol{s}$, an estimate of the discriminative ability of a binary classifier using calibration data (Colwell et al., 2014; Zweig and Campbell, 1993). In this case, a 5-fold cross-validated linear discriminant analysis was used to derive the averaged AUC (see sections 4.7.1 and 4.5.1), after a proper feature extraction process involving epoch windowing and decimation (Martínez-Cagigal et al., 2020) (see sections 4.3).

Once the objective functions are modeled, the solution associated with the global optima may be found by using a meta-heuristic; i.e., an algorithm that iteratively improves a candidate solution. Even though there are more than 70 evolutionary computation based meta-heuristics (Jr et al., 2013), only a few key strategies are repeatedly used by these algorithms and slightly modified among them. Furthermore, most of them are single-objective oriented. The aforementioned multi-objective formulation can be adapted into a single expression by using weighted approaches (Coello and Reyes-Sierra, 2006; Martínez-Cagigal and Hornero, 2017a,b; Martínez-Cagigal et al., 2018); however, the supervisor would

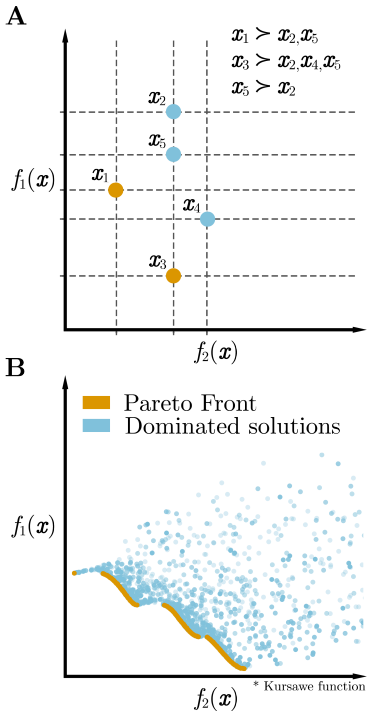
get a single solution, lacking of the possibility to consider the multi-objective optimization tradeoff. Among the well-known single-objective meta-heuristics, stand out: genetic algorithms (GA) (Holland, 1992), (binary) differential evolution (BDE, DE) (Storn and Price, 1997), (binary) particle swarm optimization (BPSO, PSO) (Kennedy and Eberhart, 1995), ant colony optimization (ACO) (Dorigo et al., 2006), or bees algorithms (BA) (Pham et al., 2006), among others.

By contrast, multi-objective algorithms try to optimize more than one objective simultaneously. Since these objectives usually conflict among themselves, it is required to introduce a new metric in order to measure the quality of each solution: the *dominance*. It is said that a solution $\boldsymbol{s}_1$ dominates another solution $\boldsymbol{s}_2$ (i.e., $\boldsymbol{s}_1 \succ \boldsymbol{s}_2$) if $\forall i : f_i(\boldsymbol{s}_1) \leq f_i(\boldsymbol{s}_2)$ and $\exists j : f_j(\boldsymbol{s}_1) < f_j(\boldsymbol{s}_2)$ –in minimization problems–. Therefore, those solutions that are not dominated by any other are considered optimal. As shown in figure 4.2, the set containing optimal solutions forms the Pareto-front curve, which depicts the tradeoff between the objectives (Deb, 2005). Concerning the BCI channel selection model of equation 4.1, the Pareto-front would display configurations that use different number of channels, allowing the supervisor to freely select one of them. Common multi-objective meta-heuristics include: non-sorting genetic algorithm II (NSGA-II) (Deb et al., 2002); multi-objective PSO (MOPSO) (Reyes-Sierra and Coello, 2005); or strength pareto evolutionary algorithm 2 (SPEA2) (Zitzler et al., 2001); among others.

4.2.1 Dual-Front Sorting Genetic Algorithm

In spite of the large amount of existing meta-heuristics, all of them should be adapted to the BCI channel selection context to assure a proper performance. The peculiarities of this discrete problem make some of the inner strategies of the meta-heuristics suboptimal or even futile. For instance, continuous or crowding distances, distance soughts, transfer functions or even any repository control approach. For this reason, we proposed a novel multi-objective algorithm especially designed for the BCI channel selection problem: the dual-front sorting genetic algorithm (DFGA) (Martínez-Cagigal et al., 2020). The pseudocode of DFGA is shown in figure 4.2(C), and its five main key aspects are detailed below:

1) **Deterministic initialization.** Heuristics generally initialize the population by generating random solutions. However, the use of deterministic initialization may reduce both the inter-run variability due to stochastic effects and a large amount of computation time. Although deterministic algorithms are not likely to provide a global optimum, DFGA considers their outputs as



C

```
1:  procedure DFGA(N_c, p_c, p_m)
2:      [R_{S,F}] ← backward_elimination()
3:      [O_{S,F}, S_{S,F}] ← dual_front(R)
4:      while ¬stopCondition do
5:          /* Parent selection */
6:          S_o ← binary_tournament(2N_c/3, O)
7:          S_s ← binary_tournament(N_c/3, S)
8:          P_S ← S_o ∪ S_s
9:          /* Genetic operators */
10:         P'_S ← crossover(P_S, p_c)
11:         P''_S ← mutation(P'_S, p_m)
12:         P''_F ← F(P''_S)
13:         /* Synthetic solutions */
14:         Q_S ← generate_synthetic(R_S)
15:         Q_F ← F(Q_S)
16:         /* Repository updating */
17:         R_{S,F} ← rep_update(R ∪ P'' ∪ Q)
18:         [O_{S,F}, S_{S,F}] ← dual_front(R)
19:     end while
20:     return O_{S,F}
21: end procedure
```

Figure 4.2: (A) Dominance concept in a two-objective minimization problem: (orange) non-dominated solutions, (blue) dominated solutions. (B) Optimization of a Kursawe function: (orange) solutions that belong to the Pareto-Front, (blue) suboptimal solutions. (C) Pseudocode of DFGA.

intermediate solutions. Regardless of their qualities, we hypothesized that these solutions are equivalent to those that will be eventually reached after several generations of a randomly initialized algorithm. In this study, backward elimination (BE, section 4.4.1) is used to initialize the repository. The algorithm begins with the full set of channels and sequentially removes the most irrelevant one (Jobson, 1991). The rejected channel in each step is the one that returns the minimum $f_1(\boldsymbol{s})$ value if removed from the model $\boldsymbol{s}$ (i.e., its inclusion does not contribute to improve the system's performance). The algorithm continues removing channels until the set is empty. Note that this operation will fill the repository $\mathcal{R}$ up with N_c solutions.

2) Elitism. In each generation, the repository is updated following an elitist approach. As depicted in the figure 4.3(A), for each unique value of $f_2(\boldsymbol{s})$ (i.e., for each number of channels), the repository solution that minimizes $f_1(\boldsymbol{s})$ is selected. Note that this operation is applied in the repository, which

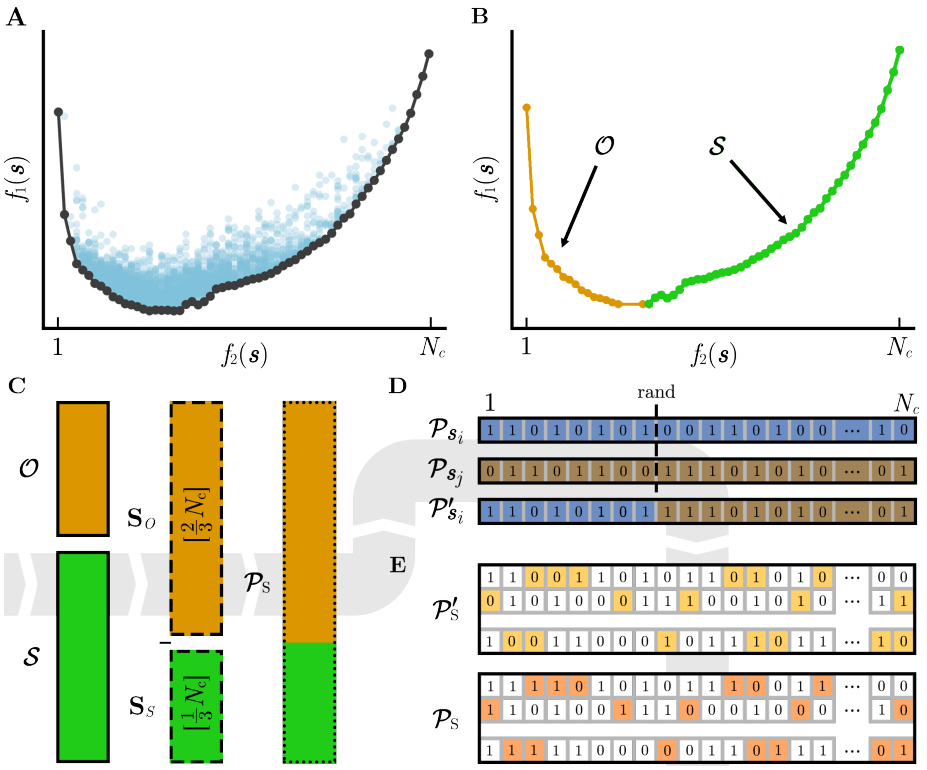


Figure 4.3: Visual aids to clarify DFGA operations: (A) elitism, (B) dual-front sorting, (C) parent selection, (D) single-point crossover and (E) bit-string mutation.

includes both non-dominated and dominated solutions, creating a balance between local and global exploitation.

3) **Dual-front sorting.** Due to the deterministic initialization, the repository should have a well-defined curve from the very beginning of the algorithm. This aspect leads to a Pareto-front that is supposed to include solutions with few channels. Traditionally, only Pareto-optimal solutions are considered in the selection stage. Despite their convenience over dominated solutions, considering only the Pareto-front will lead to a local exploitation of solutions with few channels. However, because of the intrinsic fixed size of the repository in BCI problems (i.e., limited to N_c), the exploitation of solutions with a greater number of channels was no longer an issue. Furthermore, it might favor the spreading of the Pareto-front and the global search of DFGA. According to this rationale, DFGA subdivides the repository into two sets: $\mathcal{O}$ (i.e., optimal set), which includes the non-dominated solutions; and $\mathcal{S}$ (i.e.,

sub-optimal set), which includes the dominated solutions. Dual-front sorting operation is shown in figure 4.3(B). Then, binary tournament selection is applied in both sets, selecting $^{2N_c}/_3$ solutions from $\mathcal{O}$, and $^{N_c}/_3$ solutions from $\mathcal{S}$. Note that a solution may be selected more than once in the new population. Finally, these solutions are combined in the population in order to suffer recombination (i.e., crossover) and mutation, as shown in the figure 4.3(C).

4) **Genetic operators.** Owing to the binary nature of the search space, traditional genetic operators were considered the most convenient approach for generating new solutions from a parent population. First, for each solution $\boldsymbol{x}_i$, single-point crossover is applied with probability p_c. That is, $\boldsymbol{s}_i$ and another randomly picked solution $\boldsymbol{s}_j$ $(i \neq j)$ are combined into $\boldsymbol{s}'_i \leftarrow \boldsymbol{s}_i[1:u] \cup \boldsymbol{x}_j[u+1:N]$, where $u \sim \text{rand} \in [1, N]$. For each solution, bit-string mutation is also computed with probability p_m. In other words, if the n-th bit of a solution $\boldsymbol{s}'_i$ has to be mutated, its value is flipped (i.e., $\boldsymbol{s}''_i[n] \leftarrow \neg \boldsymbol{s}'_i[n]$). The procedure is illustrated in the figure 4.3(D–E). In (Martínez-Cagigal et al., 2020), following the recommendations of the literature, we fixed p_m and p_c to $1/N_c$ and 0.90, respectively (Deb et al., 2002).

5) **Synthetic solutions.** When the values of p_c or p_m are too high, the mutated population tends to exploit the middle part of the repository. In other words, solutions with few channels tend to add more channels, whereas crowded solutions tend to decrease their number of channels. In order to maintain a similar exploitation across the entire repository spectrum, synthetic solutions are generated apart from the mutated population. However, a random generation of solutions across this spectrum will unnecessarily increase the number of evaluations, slowing down the algorithm. DFGA generates synthetic solutions using a metric that is intended to maintain the most relevant channels of the current repository. The rank of the i-th channel is defined as the number of times that the channel i is present in the repository (i.e., $r_i = |i \in \mathcal{R}|$). DFGA iteratively creates solutions that have from 1 to $N_c - 1$ channels by means of a roulette wheel selection (i.e., fitness proportionate selection) based on the rank values. It is worthy to mention that DFGA generates a total of $N_c - 1$ solutions, since the N_c-th solution that contains all the channels is already part of the repository.

4.3 Feature extraction

Owing to the fact that P300-based BCIs must work in real-time, the feature extraction stage is kept simple, being usually viewed as a signal conditioning step before applying the final pattern recognition. Fundamentally, it is composed by (1) a decimation, (2) a windowing, and (3) a concatenation across channels:

1) **Decimation.** The pre-processed EEG signal is decimated in order to reduce the dimensionality. In a nutshell, decimation is the process to reduce the sampling rate f_s to f_d by applying a low-pass filtering and down-sampling the data. Note that the low-pass filter must have a cutt-off frequency of $f_d/2$ to avoid aliasing, according to the Nyquist–Shannon theorem (Krusienski Dean J., 2012). For the online studies, we used $f_d = 20$ Hz (Martínez-Cagigal et al., 2017, 2019a) to keep up with real-time; for the channel selection study, we used $f_d = 25$ Hz (Martínez-Cagigal et al., 2020), since minimizing the amount of samples per epoch was no longer an issue because it was an offline analysis.

2) **Windowing.** Then, a time window is extracted from each epoch (i.e., flashing in RCP). Typically, it lasts from the flashing onset (i.e., $t = 0$ ms) to a timestamp delayed enough to capture the entire P300 wave (e.g., 600–1000 ms). Afterward, the epoch is normalized using a z-score of a previous reference time window; i.e., $\boldsymbol{x} = (\boldsymbol{x} - \mu_{\boldsymbol{x}_R})/\sigma_{\boldsymbol{x}_R}$, were $\boldsymbol{x}_R$ is the EEG signal of the reference window and μ and σ denote the mean and standard deviation, respectively. In the present Doctoral Thesis, epochs were extracted using a 0–800 ms time window, and a -200–0 ms reference window (Martínez-Cagigal et al., 2017, 2019a, 2020).

3) **Concatenation.** Once the previous steps are applied, the epoch has dimensions $\boldsymbol{x} \in \mathbb{R}^{N_s, N_c}$, where N_s is the number of samples of the window. In order to transform the epoch into a feature vector, channels are concatenated to get $\boldsymbol{f} \in \mathbb{R}^{1, N_s \times N_c}$.

4.4 Feature selection

Even though the decimation step reduced the dimensionality of feature vectors, the number of features is usually too large in comparison with the number of observations, which can lead to an over-fitting of the subsequent classifier. Hence, a

feature selection algorithm to remove redundant or irrelevant features is commonly applied before the pattern recognition stage (Jobson, 1991).

4.4.1 Step-Wise regression

Step-wise (SW) regression is undoubtedly the most popular feature selection algorithm in P300-based BCIs (Krusienski et al., 2006, 2008; Sellers Eric W., 2012). SW performs a sequential process, where at each step a single feature is added (i.e., *forward selection*, FS) or removed (i.e., *backward elimination*, BE) from a least square model in the next fit (Jobson, 1991).

Starting without any feature, the algorithm performs a FS by testing the significance of adding each feature separately, according to a partial F-statistic. Then, the most statistically significant feature, provided its p–value $< p_{in}$, is added to the model. After each new entry to the model, a BS procedure is performed by testing the significance again of each included feature. The less significant one, provided its p–value $> p_{out}$, is then removed from the model. The combination of FS and BS procedures is repeated until (1) there are no features that meet the p_{in} and p_{out} criteria, (2) there are no remaining features, or (3) the number of included features reach a predefined limit. It is worthy to mention that, in order to prevent an infinite loop, the p_{in} criterion should be at least as small as p_{out} (Jobson, 1991).

In the present Doctoral Thesis, SW regression was used as a feature selection stage in the online studies (Martínez-Cagigal et al., 2017, 2019a) using the following criteria: $p_{in} = 0.10$ and $p_{out} = 0.15$. Note that these criteria were set according to previous literature (Krusienski et al., 2006, 2008).

4.5 Feature classification

Once features are extracted and appropriately selected, each feature vector must be classified into two classes: presence/absence of P300, or non-/control state. Feature classification is considered a technique under the scope of the pattern recognition field, which concerns the automated recognition of regularities in data (Bishop, 2006). In turn, it is also closely related with terms such as *artificial intelligence* or *machine learning*, which are often used interchangeably.

Among the different machine learning approaches aimed at feature classification, only supervised learning techniques are in the scope of the present Doctoral Thesis. These methods infer a discriminant function from *labeled* examples (i.e.,

training data), which is further used to classify new examples (i.e., test data) into different classes. Therefore, before using a P300-based BCI system online, the classifier must be calibrated in a previous copy-spelling session. In that session, users are asked to spell a predefined word via RCP or CS, labeling the flashings that contains the target character as positive; otherwise as negative. The classifier is trained with those examples, fitting a discriminant function that sets a hyperplane that separates both classes. Afterward, the new real-time data of the online sessions, which is not labeled anymore, is continuously classified by the model. Note that, for the asynchronous systems, also labeled non-/control examples were recorded as a calibration step (Martínez-Cagigal et al., 2017, 2019a,b).

4.5.1 Linear Discriminant Analysis

Linear discriminant analysis (LDA) is a method that finds a linear combination of features that try to separate two or more classes. Since the data is projected down to a lower dimension, LDA is not only used to classify events, but also as a dimensionality reduction technique (Bishop, 2006). For all linear classifiers, the projection is expressed as $y(\boldsymbol{X}) = \boldsymbol{w}^T\boldsymbol{X}$, where $\boldsymbol{X} \in \mathbb{R}^{n,m}$ with n features and m observations, and $\boldsymbol{w}$ is the weight vector. In LDA, the projection will try to (1) maximize the distance of the mean of both classes, and simultaneously (2) minimize the variance of each class; the so-called Fisher criterion. Indeed, LDA is viewed as a generalization of the Fisher's linear discriminant (FLD), whose solutions are equivalent in two-class problems. Furthermore, neither LDA or FLD make assumptions about normal distributions in two-class problems (Pelillo, 2013), making them the most popular classification algorithms in P300-based BCIs (Krusienski et al., 2006, 2008; Sellers Eric W., 2012).

For a two-class problem like the one that concerns us, the weight vector can be identified by solving the following optimization problem:

$$\max J(\boldsymbol{w}) = \frac{\boldsymbol{w}^T\boldsymbol{S}_B\boldsymbol{w}}{\boldsymbol{w}^T\boldsymbol{S}_W\boldsymbol{w}}, \tag{4.2}$$

$$\boldsymbol{S}_B = (\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)^T, \tag{4.3}$$

$$\boldsymbol{S}_W = \sum_{n\in\mathcal{C}_1}(\boldsymbol{X}_n - \boldsymbol{\mu}_1)(\boldsymbol{X}_n - \boldsymbol{\mu}_1)^T + \sum_{n\in\mathcal{C}_2}(\boldsymbol{X}_n - \boldsymbol{\mu}_2)(\boldsymbol{X}_n - \boldsymbol{\mu}_2)^T, \tag{4.4}$$

where $\boldsymbol{S}_B$ indicates the between-class matrix, $\boldsymbol{\mu_i}$ the mean value of the features of classes $i \in \mathcal{C}_{\{1,2\}}$, and $\boldsymbol{S}_W$ the within-class matrix. The optimal weight vector

would be:

$$\boldsymbol{w} \propto \boldsymbol{S}_W^{-1}(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1). \tag{4.5}$$

Therefore, a new observation $\boldsymbol{x}$ is classified as $\mathcal{C}_1$ if $y(\boldsymbol{x}) \geq y_0$; otherwise is classified as $\mathcal{C}_2$ (Bishop, 2006). Generally, the threshold $y(\boldsymbol{x}) = y_0$ is chosen as the hyperplane between the projections of both means: $y_0 = \boldsymbol{w}^T \cdot \frac{1}{2}(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2)$. However, in P300-based BCIs that use visual oddball paradigms such as RCP, CS or RSVP, an additional step must be performed to determine the selected command in real-time. A categorical classification is not needed, but a *score* that indicates the likelihood to belong to the positive class. Generally, the score is the posterior probability, computed by modeling the predicted training examples of the positive class as multivariate gaussian distribution and estimating its mean and covariance through a max-likelihood approach; i.e., $l = \hat{P}(\mathcal{C}_1|y(\boldsymbol{x})) \sim \mathcal{N}(\boldsymbol{\mu}_1, \Sigma_1)$ (Bishop, 2006).

Hence, the scores $\boldsymbol{l}$ of each flashing group (i.e., rows and columns in RCP, shapes in CS, individual symbols in RSVP) are averaged. In RSVP/CS, the selected symbol/shape would be the one that yields the maximum score, l_{max}; whereas in RCP, the selected command would be the intersection between the selected row and column. Similarly, the selected row/column would be the row/column that yields the maximum score.

4.6 Asynchrony management

As mentioned in sections 1.4.4 and 1.6, it is required to detect the non-/control states in real-time in order to reach a comprehensive asynchronous control in visual P300-based BCIs. Thus, an asynchrony management could be viewed as a supplementary signal processing pipeline independent of the P300 detection. In this section, two proposed asynchronous methodologies are presented: the thresholding and the multiscale entropy approaches.

4.6.1 Thresholding

The thresholding approach is based on the hypothesis that control state ERPs yield higher l_{max} scores than non-control ones. The rationale behind this lies in the fact that P300 should not be elicited when users do not attend to the stimuli, causing lower posterior probabilities of the positive class. Therefore, the objective is to set a constant threshold that can distinguish between control and non-control l_{max} values.

Once enough training examples of non-/control states are recorded, scores are concatenated and labeled according to their class: $\boldsymbol{l}_C$ contains control scores; whereas $\boldsymbol{l}_N$ contains non-control ones. These vectors are fed into a ROC curve, which depicts the performance of a binary discriminant when the threshold varies along the input values (Zweig and Campbell, 1993). The curve displays the true positive rate (i.e., sensitivity) against the false positive rate (i.e., $1 -$ sensitivity). The optimal threshold τ would be the one that maximizes the sensitivity-specificity pair; i.e., the point that yields the minimum Euclidean distance to coordinates (0,1). Therefore, T is estimated in a calibration phase with training data. Whenever the system classifies a new trial (i.e., character) in an online setup, the l_{max} score is compared with the estimated threshold: if $l_{max} \geq T$, the trial is classified as a control state; otherwise, it is considered non-control. Finally, control state trials are delivered, whereas non-control ones are ignored. This method was proposed and tested in the web browser study (Martínez-Cagigal et al., 2017), and afterward applied in the social networks application study (Martínez-Cagigal et al., 2019a).

4.6.2 Multiscale sample entropy

In Martínez-Cagigal et al. (2019b), we proposed the multiscale entropy (MSE) based on sample entropy (SampEn) as a feature extraction metric that could characterize non-/control states. MSE is a nonlinear algorithm that estimates the complexity of a signal by assessing entropy changes in different time scales (Costa et al., 2002). Information about the dynamical structure is provided through sequential entropy measurements in coarse-grained versions of the original signal, considering a signal to be more complex if reaches higher entropy values for most scales (Costa et al., 2005). In MSE, the τ-th scaled coarse-grained signal should be obtained by decimating the original one by a factor of τ (see section 4.3) (Humeau-Heurtier, 2015). Hence, the algorithm computes an entropy metric from the original time series ($\tau = 1$) to the highest considered scale ($\tau = \tau_{max}$).

SampEn is a nonlinear method that estimates the irregularity of one-dimensional physiological signals by assigning higher values to those that show larger degrees of disorder (Richman and Moorman, 2000). Inside a BCI context, SampEn is applied for each channel separately, returning a single feature for each of them. In a nutshell, the algorithm measures the conditional probability that a template of m consecutive samples, which already matches another template, still

matches it if their lengths are increased by one sample (Humeau-Heurtier, 2015):

$$\text{SampEn}(m, r, N) = \lim_{N \to \infty} -\ln\left(\frac{A^m(r, N)}{B^m(r, N)}\right), \tag{4.6}$$

where m denotes the embedding dimension, r the tolerance, N the signal length, and $B^m(r, N)$ and $A^m(r, N)$ denote the probabilities of matching for templates of m and $m + 1$ points, respectively. Considering a signal $\boldsymbol{x} = [x_1, x_2, \ldots, x_N]$, a template vector of length m would be $\boldsymbol{x}_m(i) = [x_i, x_{i+1}, \ldots, x_{i+m-1}]$. In practice, a match between two templates $\boldsymbol{x}_m(i)$ and $\boldsymbol{x}_m(j)$ occurs provided that $d[\boldsymbol{x}_m(i), \boldsymbol{x}_m(j)] < R$; where $d[\cdot]$ denotes the Chebyshev distance and the tolerance value R is usually dependent of the standard deviation of $\boldsymbol{x}$ (i.e., $R = r \cdot \sigma_x$) (Humeau-Heurtier, 2015; Richman and Moorman, 2000). Therefore, the SampEn is estimated as follows:

$$\text{SampEn}(m, r, N) = -\ln\left(\frac{N - m + 1}{N - m - 1} \cdot \frac{A}{B}\right), \tag{4.7}$$

where B and A denote the number of templates of lengths m and $m + 1$ that matches for each different combination of i and j (given $i \neq j$), respectively. The estimator is unbiased due to the normalization; however, its variance decreases when the length of the signal increases. For this reason, the estimation is considered accurate if $N \geq 10^m$ (Humeau-Heurtier, 2015; Richman and Moorman, 2000).

The application of MSE to characterize asynchronous states in BCIs should guarantee that signals are large enough to provide an accurate estimator, requiring a slightly different feature extraction procedure. In this case, non-/control training trials (i.e., characters) were decimated (see section 4.3), but epochs were extracted according to the figure 4.4. Therefore, the i-th epoch would be the decimated signal from the first onset of the current character to the end of the i-th sequence. MSE was then applied to each epoch, channel and subject. Note that only the scales that provided coarse-grained signal whose length $N \geq 10^m$ were computed. Hyperparameters were optimized inside the common range of physiological signals: $m \in \{1, 2\}$, $r \in (0.10, 0.15, \ldots, 0.30)$, and $\tau = (1, 2, \ldots, 25)$; under a leave one subject out procedure (see section 4.7.1) (Martínez-Cagigal et al., 2019b).

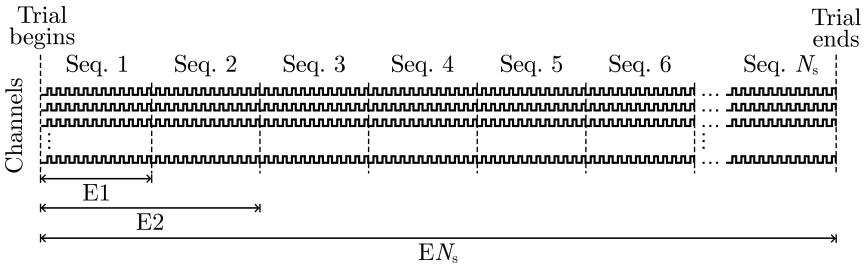


Figure 4.4: Epoch extraction procedure of a single trial for the application of SampEn-based MSE. The number of possible epochs (E) matches the total number of sequences.

4.7 Performance assessment

During the different studies that compose the present Doctoral Thesis, several validation methods were applied to assess the performance of the proposed signal processing methods and applications.

4.7.1 Cross-validation

Fundamentally, cross-validation (CV) is a model validation technique that assesses the generalization ability of the results of a statistical analysis. In other words, CV allows to evaluate if a model can generalize to an independent dataset; which is not assured if training samples were not representative (Witten and Frank, 2011).

1) **Folds.** In a k-fold CV, the training dataset is split into k equal partitions. Then, the algorithm sequentially uses one of the folds for testing, keeping the remainder dataset to train the model. That is, every instance has been used for testing exactly once, returning a diagnostic ability statistic (e.g., accuracy) for each fold. Therefore, providing an adequate estimate of the generalization ability of the model when averaging these statistics (Witten and Frank, 2011). This algorithm was used in Martínez-Cagigal et al. (2020) to estimate the performance of the single- and multi-objective algorithms for channel selection.

2) **Leave one out (LOO).** The LOO CV is a particularization of the k-fold CV that keeps only one sample (or subject) for testing (i.e., equivalent to an m-fold CV if the training dataset is composed by m observations) (Witten and Frank, 2011). This technique was used in Martínez-Cagigal et al. (2019b) to optimize the MSE hyperparameters and to estimate the performance of the signal processing pipeline to reach an asynchronous control of the BCI.

4.7.2 Speed rates

When a practical P300-based BCI application to assist the motor-disabled is developed, not only the performance should be evaluated, but also the speed of the system for selecting commands.

1) **Information transfer rate (ITR).** Traditionally, ITR has been employed as a metric to measure the speed of BCI systems. Based on Shannon's information theory, the ITR estimates the number of bits per trial selection:

$$\text{ITR} = \frac{1}{T}\left(\log_2 N + P\log_2 P + (1-P)\log_2\frac{1-P}{N-1}\right), \tag{4.8}$$

where N is the number of trial selections, P is the selection accuracy, and T_s is the average duration of a trial in seconds (Wolpaw et al., 2000). However, the ITR assumes that: (1) the system is memoryless, (2) all possible characters are equally probable, and (3) a synchronous paradigm is used. The metric is thus biased for asynchronous BCIs because it neither considers multiple character matrices, nor correction of erroneous selections, nor asynchronous pauses (Speier et al., 2013).

2) **Output characters per minute (OCM).** In contrast to ITR, OCM is a communication rate metric especially suitable for asynchronous systems. The metric is simply the ratio of the number of selections to the duration of the task in minutes; i.e., OCM $= N/T_{min}$ (Speier et al., 2013). OCM was used in Martínez-Cagigal et al. (2019a) to evaluate the speed of the developed social networking BCI system.

4.8 Statistical analysis

In order to make a fair comparison between two groups of results, several statistical tests were applied in the compendium of publications. For instance, to assess significant differences between non-/control features (Martínez-Cagigal et al., 2019b), or between results obtained from controls and patients (Martínez-Cagigal et al., 2017, 2019a).

4.8.1 Hypothesis testing

A statistical test evaluates the evidence the data provides against the null hypothesis H_0, which usually states that data is generated by random processes. The

returned p-value indicates the probability of getting an fluctuation away from the H_0 distribution, assuming that H_0 is true. It is considered that H_0 is false (i.e., that there are *significant differences* in the data) if p-value $< \alpha$, where $\alpha = 0.05$ is the level of significance (Narsky and Porter, 2013).

In the present Doctoral Thesis, two different univariate, continuous, non-parametric statistical tests have been used. Before their application, it was stated that data were not normal and homoscedastic. When the comparisons were paired (i.e., dependent, when data came from the same subjects), Wilcoxon signed-rank tests were used. For instance, comparisons such as non-/control features or changes in measurements between sessions (Martínez-Cagigal et al., 2017, 2019a,b). For unpaired comparisons (i.e., independent, when data came from different populations), such as the ones that compare control subjects versus patients, Mann–Whitney U tests (i.e., Wilcoxon rank-sum tests) were applied (Martínez-Cagigal et al., 2017, 2019a).

4.8.2 Multiple testing correction

In the literature, the result of a test is considered significant if p-value $< \alpha$; i.e., the probability of the comparison not being significant (false positive, type I error) is less than α. However, when N independent tests are applied, the probability of making at least one type I error by chance drastically increases to $1 - (1 - \alpha)^N$ (e.g., 99.41% for $N = 100$ and $\alpha = 0.05$), the so-called family-wise error rate (FWER). This phenomenon is known as the *multiple comparisons problem* or the *look-elsewhere effect*, and requires to establish a stricter α value or to adjust the p-values to compensate the inferences being made (Narsky and Porter, 2013).

Multiple testing can be addressed by correcting the FWER (e.g., Bonferroni, Šidák, Holm-Bonferroni) or the false discovery rate (FDR) (e.g., Benjamini-Hochberg, Benjamini-Yekutieli, Storey q-values) (Farcomeni, 2008). The latter is defined as the expected proportion of type I errors among the tests that have already been considered significant. In this Doctoral Thesis, the procedure of Benjamini-Hochberg (BH) was applied to correct the p-values by assuring that FDR $= \alpha$ (Martínez-Cagigal et al., 2019b). The algorithm (1) sorts the obtained p-values in descending order, $\boldsymbol{p}$; (2) calculates $\hat{\boldsymbol{p}} = Np_k/k$ for $k = 1, 2, \ldots, N$; and finally (3) adjusts the p-values by computing the cumulative minimum of vector $\hat{\boldsymbol{p}}$, i.e. $\boldsymbol{p}_k^* = \min \hat{p}_{1,\ldots,k}$ (Benjamini and Hochberg, 1995).

Chapter 5

Results

In this chapter, the most relevant results of the Doctoral Thesis are summarized. They are organized according to the hypotheses of section 2.1, which in turn have almost a directly relation with the papers that encompasses the compendium of publications (see appendix A).

5.1 Asynchronous management using thresholding and entropy metrics

Asynchronous management in P300-based BCI systems was assessed by two different approaches: thresholding (section 4.6.1), and MSE (section 4.6.2).

The thresholding approach was implemented in the assistive applications. Its viability was tested online with motor-disabled subjects (Martínez-Cagigal et al., 2017, 2019a) according to the procedure detailed in section 4.6.1. Figure 5.1 shows the normalized SW-LDA scores of the selected characters for non-/control calibration runs in both applications. As shown, the discriminative ability of the threshold is higher for HS, reaching averaged training accuracies of 96.00%±4.77% and 96.74%±3.47% for the web browser and social networking applications, respectively. By contrast, these accuracies were reduced for the patients, yielding 86.77%±7.47% and 84.31%±9.16%, respectively. The predictive validity of the thresholds was thus expected to decay in the evaluation sessions. However, it should be noted that the purposes of both papers were oriented toward the evaluation of the web browser and social networking applications. Therefore, the asynchronous management was not isolated and online performances were evalu-

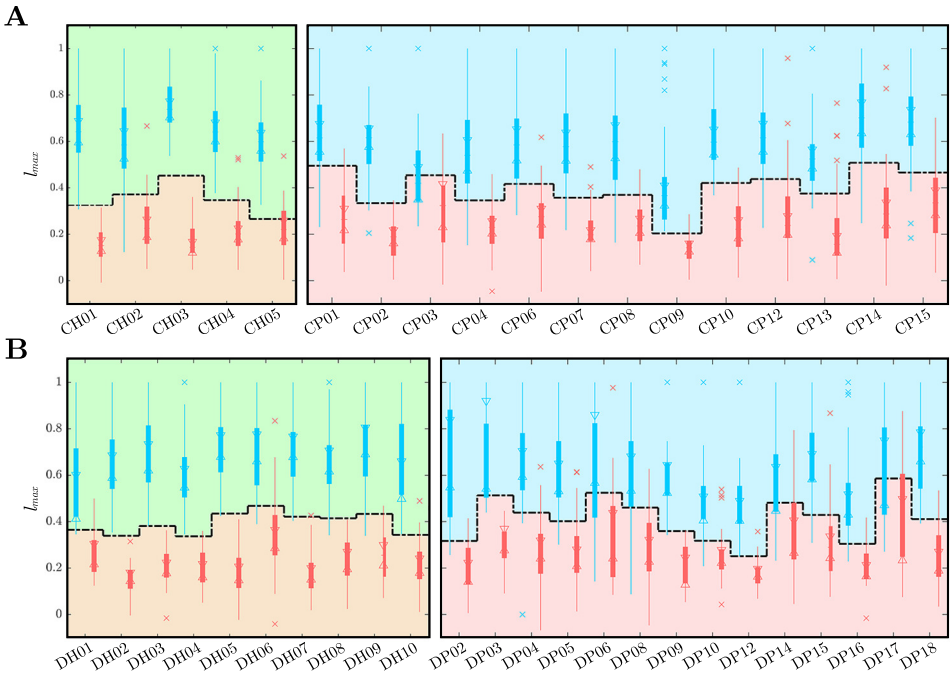


Figure 5.1: Asynchronous threshold estimation in the calibration sessions of (A) web browser and (B) social networking applications. Boxplots depict normalized SW-LDA scores of the selected characters for control (blue) and non-control (red) runs. The averaged threshold value of each user is also shown as a dash-dotted line.

ated using thresholding as a part of the overall system (see sections 5.3 and 5.4). In this regard, the false negative rate (FNR), defined as the ratio of false negatives to the total number of selections, is an excellent measure to estimate the benefit from using an asynchronous approach. HS reached an averaged FNR of 4.61%±6.48% and 1.10%±3.17% for the web browser and social networking applications, respectively; whereas motor-disabled patients obtained 10.87%±12.46% and 1.54%±4.18%.

Non-/control state signals were also characterized offline by applying SampEn-based MSE, according to the procedure described in section 4.6.2. As aforementioned, EEG recordings of 10 HS attending (i.e., control) and ignoring (i.e., non-control) the RCP stimuli were recording with a 16-channel cap (see section 3.1). Then, data was randomly divided into optimization (36 trials per user) and validation (84 trials per user) sets. Hyperparameters were optimized using a wrapper LDA-based leave one subject out procedure (see sections 4.5.1 and 4.7.1) over the optimization set (Martínez-Cagigal et al., 2019b).

Table 5.1: Validation results for each user in the asynchronous SampEn-based MSE approach.

	N_s	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Accuracies (%)	**AH01**	71.4	77.4	81.0	85.7	88.1	88.1	89.3	90.5	94.1	94.1	92.9	92.9	94.1	95.2	94.1
	AH02	83.3	88.1	89.3	85.7	89.3	89.3	91.7	91.7	91.7	90.5	92.9	91.7	92.9	94.1	92.9
	AH03	83.3	82.1	88.1	83.3	86.9	90.5	88.1	90.5	94.1	92.9	92.9	92.9	92.9	92.9	92.9
	AH04	61.9	78.6	81.0	75.0	75.0	75.0	81.0	81.0	79.8	81.0	83.3	90.5	89.3	91.7	91.7
	AH05	72.6	70.2	72.6	78.6	78.6	82.1	89.3	89.3	91.7	91.7	91.7	94.1	95.2	96.4	96.4
	AH06	89.3	94.1	96.4	96.4	95.2	94.1	96.4	95.2	94.1	95.2	96.4	96.4	96.4	97.6	98.8
	AH07	75.0	89.3	92.9	95.2	96.4	96.4	95.2	95.2	92.9	94.1	95.2	95.2	96.4	96.4	95.2
	AH08	77.9	80.6	85.7	86.9	86.9	88.1	86.9	84.5	89.3	89.3	86.9	89.3	90.5	89.3	89.3
	AH09	78.6	90.5	91.7	88.1	94.1	90.5	92.9	95.2	95.2	92.9	95.2	95.2	97.6	95.2	96.4
	AH10	76.2	86.9	91.7	95.2	95.2	92.9	94.1	92.9	95.2	97.6	97.6	96.4	96.4	96.4	96.4
	Mean	**76.9**	**83.8**	**87.0**	**87.0**	**88.6**	**88.7**	**90.5**	**90.6**	**91.8**	**91.9**	**92.5**	**93.5**	**94.2**	**94.5**	**94.4**
	SD	7.6	7.2	7.1	7.1	7.2	6.2	4.6	4.7	4.6	4.5	4.4	2.5	2.8	2.6	2.8
CC	**Mean**	0.8	3.5	8.2	14.6	22.6	32.5	43.7	54.9	69.7	86.3	104.9	125.2	146.4	170.6	196.8
	SD	1.0	0.3	0.8	1.0	1.4	2.0	3.1	3.3	3.8	4.7	5.6	6.0	6.5	7.2	8.6

N_s: number of sequences; CC: computational cost (in ms), SD: standard deviation. These results were obtained using the optimal hyperparameters: $\tau = 2$, $m = 1$, $r = 0.3$.

The parameters that maximized the accuracy over the optimization set were $m = 1$, $r = 0.3$ and $\tau = 2$. It was observed that accuracies show a decreasing tendency as the scale increased, regardless of the embedding dimension. Furthermore, the variation of the tolerance did not make a significant impact on the results, especially for the smallest scales. Figure 5.2(A) depicts the grand-averaged MSE values from the optimization set among all subjects for control (blue) and non-control (red) trials. Note that differences between both states were also statistically significant (p-values < 0.05, Wilcoxon signed-rank test, FDR-BH corrected) for all scales and channels. At the subject level with the fixed hyperparameters (see Figure 5.2(B)), these differences were mainly significant over the prefrontal and occipital lobes.

Table 5.1 summarizes the results of the validation stage. An average accuracy of $94.49\% \pm 2.81\%$ in classification was reached using 15 sequences. As expected, the higher the number of sequences, the higher performance. Since the optimal scale was found to be $\tau = 2$, an online setup would only compute the SampEn over the decimated signal, ignoring the remaining scales. In order to demonstrate the viability of the proposed framework, an average of 1000 iterations of the algorithm were computed to estimate the computational cost under an Intel Core i7-7700 CPU @ 3.60GHz (32 GB RAM, Windows 10, MATLAB®2018a). As shown, the averaged elapsed time when using 15 sequences was 196.78 ± 8.64 ms.

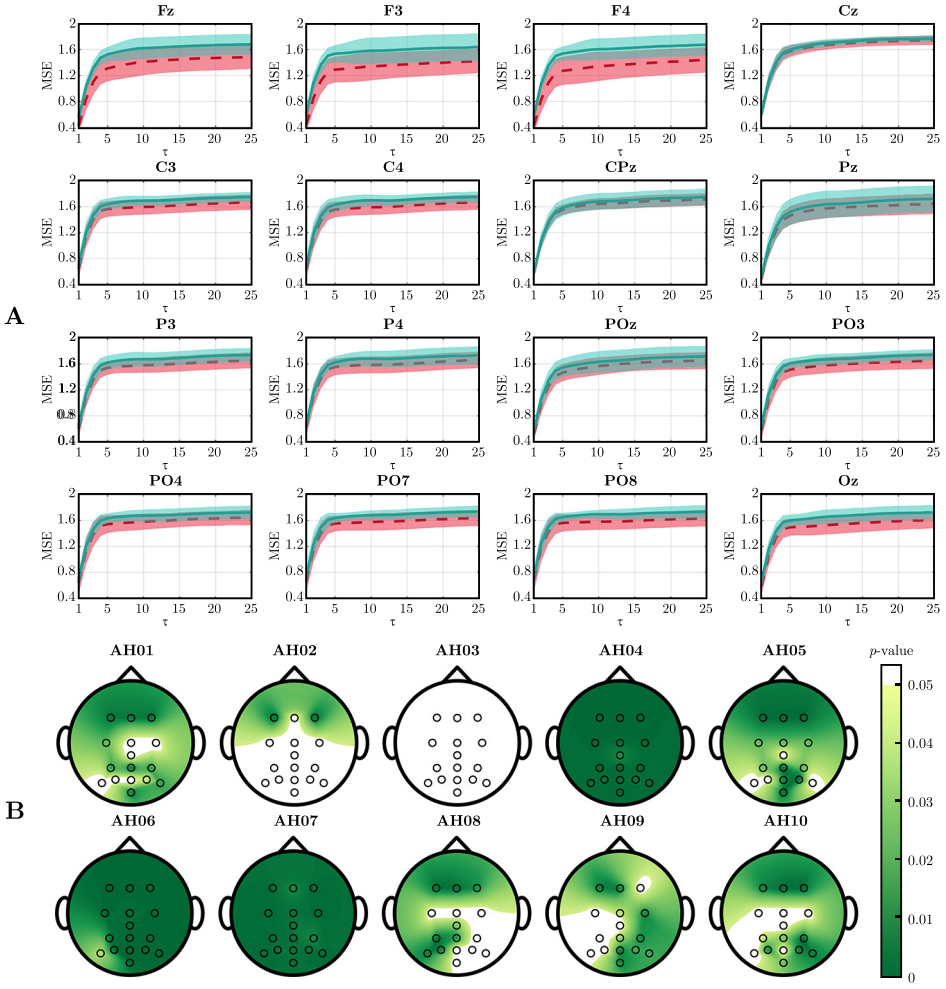


Figure 5.2: (A) Grand-averaged MSE values from the optimization set of control (blue) and non-control (red) states. Embedding dimension and tolerance were set to $m = 1$ and $r = 0.3$, respectively. Solid lines indicate the average values, whereas shaded areas indicate the standard deviation across observations. (B) Significant differences between control and non-control features from the optimization set of each user. Hyperparameters were set to the optimal values ($\tau = 2$, $m = 1$, $r = 0.3$). Significant results are colored (p-value < 0.05, Wilcoxon signed-rank test, FDR-BH corrected), whereas non-significant results are depicted in white.

5.2 Novel meta-heuristics as channel selection methods

The usefulness of meta-heuristics as channel selection methods for P300-based BCIs was explored in (Martínez-Cagigal et al., 2020). A total of 3 single- (GA, BDE, BPSO) and 3 multi-objective (NSGA-II, BMOPSO, SPEA2) existing evolutionary algorithms were compared with three public databases that used RCP, CS and RSVP paradigms (see sections 3.1 and 3.2). Owing to the limitations of these meta-heuristics when applied to this discrete problem, DFGA was also proposed as a novel multi-objective algorithm especially designed to the BCI framework. In order to assure a fair comparison among the algorithms, all of them used $m = 20$ individuals and performed a total of 4000 evaluations. Furthermore, all the algorithms were computed 20 times in order to avoid local minima .

Convergence analysis of the meta-heuristics, in function of the database, is summarized in figure 5.3. Averaged convergence curves of single-objective algorithms show the evolution of the aggregated objective function $F(\boldsymbol{s})$ across generations, estimating the ability of each method to search for an optimal solution in training phase. Training pareto-optimal solutions of the multi-objective approaches across subjects are also depicted. Figure 5.4 shows the grand-averaged ranks of selected channels for each algorithm; i.e., a rank is the normalized number of times that a channel was selected across the repetitions.

Results of the testing phase of single- and multi-objective meta-heuristics are detailed in table 5.2 and figure 5.5, respectively. Table 5.2 details the averaged testing accuracies of the solution that reached the minimum $F(\boldsymbol{s})$ value across the repetitions. The increase in accuracy of GA, BDE and BPSO in comparison with the entire set of channels and the Krusienski's set was statistically significant for almost all subjects (i.e., p-values<0.05, Wilcoxon signed-rank test, FDR-BH corrected). Differences among GA, BDE and BPSO were not significant. Figure 5.5, by contrast, shows the testing accuracies of those solutions that formed the Pareto-front, using the maximum number of sequences available.

A computational analysis was also performed under an Intel Core i7-7700 CPU @ 3.60 GHz, 32GB RAM, Windows 10 Pro, using MATLAB® 2018b. In order to estimate the computational burden of each algorithm, the required time to evaluate a solution was measured: NSGA-II (331 ms), DFGA (591 ms), GA (785 ms), BDE (810 ms), SPEA2 (835 ms), BMOPSO (852 ms), and BPSO (858 ms). From a complexity point of view, all the multi-objective meta-heuristics, including DFGA, behave as $O(N_o m^2)$, where N_o is the number of objectives and m is the

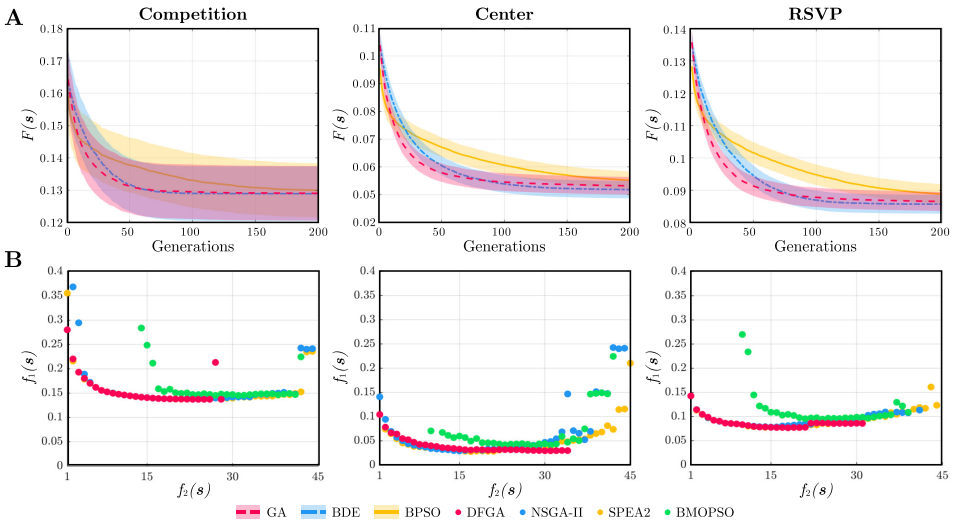


Figure 5.3: Convergence analysis of the applied meta-heuristics in the channel selection study for each database. (A) Grand-averaged convergence curves of single-objective meta-heuristics (GA, BDE, BPSO) in function of the aggregated $F(\boldsymbol{s})$ objective. Mean values are displayed with solid lines, whereas the 95% confidence interval of the subjects' repetitions is indicated by the shaded area. (B) Pareto-optimal solutions of multi-objective approaches (DFGA, NSGA-II, SPEA2, BMOPSO) across subjects and repetitions.

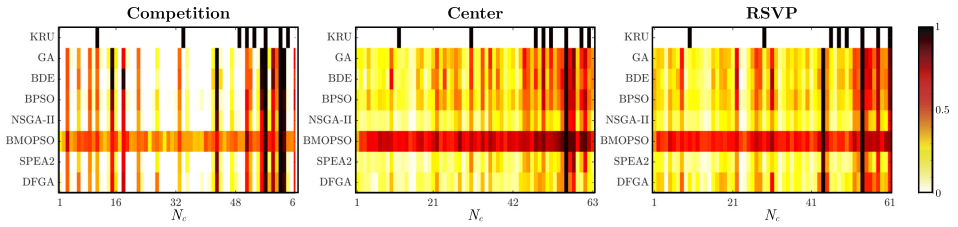


Figure 5.4: Grand-averaged channel ranks of each meta-heuristic (single: GA, BDE, BPSO; multi: NSGA-II, BMOPSO, SPEA2, DFGA) in function of the database. Krusienski's set (KRU) is also depicted.

population size.

5.3 An asynchronous assistive application for web browsing

The application of the asynchronous thresholding method led to the development of a practical assistive P300-based BCI application for web browsing. The sys-

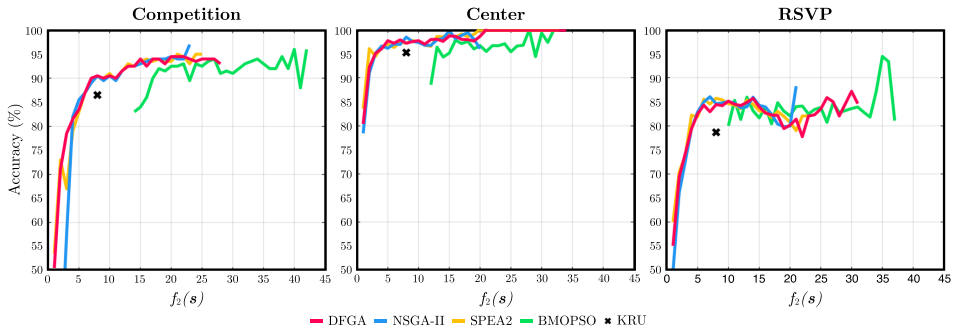


Figure 5.5: Testing character accuracies of the final Pareto-fronts returned by multi-objective meta-heuristics (DFGA, NSGA-II, SPEA2, BMOPSO) for the averaged subjects of each database. For comparison purposes, Krusienski's set (KRU) is also depicted as a cross.

Table 5.2: Averaged testing accuracies and number of channels across users of the selected run for each single-objective meta-heuristic.

Mtd.	Competition		Center		RSVP	
	Acc.	***N***	**Acc.**	***N***	**Acc.**	***N***
GA	92.0%	14.0	97.4%	12.4	84.6%	13.4
BDE	92.0%	14.5	97.9%	12.5	85.5%	13.4
BPSO	92.0%	14.0	96.8%	12.5	85.0%	13.7
ALL	92.0%	64.0	86.5%	63.0	80.3%	61.0
KRU	86.5%	8.0	95.2%	8.0	78.6%	8.0

Mtd.: method, Acc.: accuracy, N: no. of sequences. Results obtained using the maximum number of sequences available for each database (competition: 15, center: 10, RSVP: 10).

tem was composed by three stages: (1) data acquisition, which recorded and pre-processed the EEG signals; (2) processing, which encompassed the feature extraction and classification, as well as the asynchronous management; and (3) web surfing, which translated the selections into browser commands and returned feedback to the users (Martínez-Cagigal et al., 2017).

Figure 5.6 depicts the graphical user interface. The application displayed the Google Chrome web browser on the left side of the screen, reserving space for the RCP matrices on the right side. In particular, two different RCP matrices were used due to the great amount of commands required to perform a comprehensive web surfing: (i) navigation and (ii) keyboard matrices. The small size of the former (5×3) allowed users to select the commands quickly and perform a smooth navigation. The latter (9×5), by contrast, included alphanumeric characters and symbols, because it was intended to write e-mails and fill out forms. Note that the user could toggle between both by selecting the command `MTX`. Signal processing,

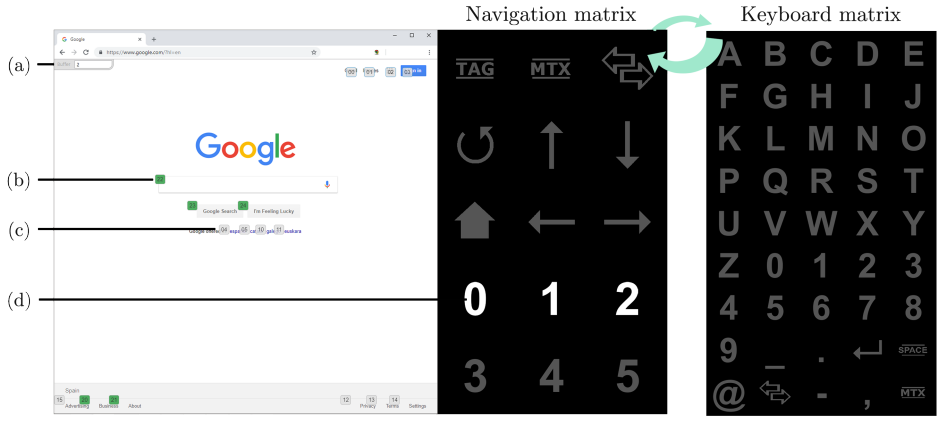


Figure 5.6: Graphical user interface of the asynchronous P300-based BCI web browser. As indicated by the buffer (a), the user introduced a '2', highlighting the potential selections (b) in green (i.e., codes that start with '2'). The rest of nodes (c) remain in gray. In this example, the navigation matrix is active and the third row (f) is currently flashing. However, users can freely toggle between both matrices by selecting the command `MTX`.

as well as stimuli generation, were coded in C++ under the BCI2000 open source platform (Schalk et al., 2004). Web surfing was based on a labeling strategy. Through a JavaScript extension, the web browser tagged each node (i.e., any clickable object) with a numeric coding, making it possible to simulate the click over them when the user selected the coding via RCP. The selection speed was prioritized by tagging the nodes using numbers from 0 to 5 (i.e., those included in the navigation matrix), and by instantaneously executing the link provided the user already selected the required number of characters, avoiding the need of an extra confirmation. Feedback was provided in several ways. On the one hand, as soon as the tags were displayed, a buffer that indicates the past selections appeared on the upper left corner of the screen. Users could remove the last selection using ←. On the other hand, potential selections (i.e., nodes whose coding starts with the previously selected characters) are highlighted in green, whereas normal tags are colored in gray.

The viability and efficacy of the assistive application was assessed by 16 MS patients and 5 HS (see section 3.1 for further details). Each user carried out four different sessions:

1) **Cal-I.** The first calibration session was intended to train the classifier and calibrate the asynchronous threshold. SW-based LDA training, composed by 4 copy-spelling runs of 6 trials (15 sequences, keyboard matrix), lasted

approximately 24 min. Threshold calibration was made of 8 runs of 6 trials, of which half were control and the other half non-control. In this stage, the navigation matrix and the optimal number of sequences of each user were used in order to reduce the duration of the calibration. In control runs, users were asked for paying attention to the requested items; whereas in non-control runs, users were asked for reading an external text while ignoring the RCP stimuli.

2) **Cal-II.** The second calibration session was only composed by another threshold calibration block. This step was designed to increase the robustness of the asynchronous management against the inter-session variability of the EEG (Picton, 1992). The final threshold value was computed as the average of the optimal thresholds of Cal-I and Cal-II.

3) **Eval-I.** The first evaluation session was composed of a series of tasks that required a comprehensive use of the web browser. In the four first ones, the asynchrony management was disabled; while the fifth was designed to test the effectiveness of the threshold. All of them are detailed in the table 5.3, with the optimal number of selections and the matrices that were required to accomplish them. Note that, if the users made a mistake, they were asked for solving it.

4) **Eval-II.** The second evaluation session assessed three additional tasks with the threshold enabled. These tasks are also detailed in table 5.3. At the end, users were asked for fulfilling a 7-point Likert questionnaire of 20 items to collect their impressions and suggestions. Positive and negative items were alternated to deal with the acquiescence bias.

The average duration to accomplish each task for both HS and MS patients is also shown in table 5.3. Furthermore, reached accuracies of each subject and sessions, as well as the optimal number of sequences and OCM, are detailed in table 5.4. Note that the accuracy is defined as the ratio of the amount of correct delivered selections to the amount of total selections, including false negatives of the threshold (i.e., those that were not delivered because they were wrongly considered as non-control selections). As shown, MS patients that could not reach a minimum accuracy of 70% in any calibration session were discarded from the subsequent assessment. Results of the questionnaires are shown in figure 5.7. As shown, almost all negative items were rated behind the neutral value, and all positive items were rated above it. Exceptions indicate that MS patients thought

Table 5.3: Evaluation tasks for the assistive web browser application.

	Task	AT	Description	OS	RM	AD	
						HS	MS
Eval-I	Link selection	No	Scroll up and down a Wikipedia webpage and select a link	6	N	2:33	4:01
	Google searching	No	Select the search form, introduce 'BCI' and select '↵'	8	N,K	4:28	6:00
	Tweeting	No	Select the Twitter form, write 2 characters and send the *tweet*	6	N	2:38	4:13
	Writing an e-mail	No	Read an inbox e-mail and reply it	13	N,K	6:18	8:18
	Passive reading	Yes	Read a piece of news and ignore the RCP	10	N	4:17	5:17
Eval-II	Reading and link selection	Yes	Scroll a Wikipedia webpage, read the information and select a link	8	N	4:18	4:44
	Tweeting	Yes	Same as Eval-I	6	N	3:25	3:44
	Active reading	Yes	Read a piece of news, scroll down if needed	4	N	1:58	2:20

AT: asynchronous threshold enabled?, OS: optimal selections, RM: required matrices (N: navigation matrix, K: keyboard matrix), AD: average duration in minutes:seconds, HS: healthy subjects, MS: multiple sclerosis patients.

that it took much too long to surf the Internet with the system, and that both MS and HS were slightly happy that the assessment sessions were over. HS also reported that they could not image themselves using the system in their daily life, which was expected due to their lack of motor disabilities. As suggestions, MS wondered whether it could be possible to increase command selection speed, plan shorter sessions, make flashings less annoying or add a `TAB` command. HS suggested to increment the number of symbols and reported that sometimes they unintentionally focused on adjacent commands.

Table 5.4: Accuracy results of the assessment sessions for the assistive web browser application.

	Subjects	Cal-I	Cal-II	N_s	Eval-I					Average[1]		Eval-II			Average	
					T1.1	T1.2	T1.3	T1.4	T1.5	Acc.	OCM	T2.1	T2.2	T2.3	Acc.	OCM[2]
HS	CH01	100.0%	100.0%	7	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	2.77	100.0%	83.33%	100.0%	94.44%	2.67
	CH02	100.0%	91.67%	11	100.00%	75.00%	100.0%	75.00%	100.0%	82.50%	1.89	92.31%	100.0%	60.00%	89.29%	1.85
	CH03	100.0%	100.0%	6	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	2.99	100.0%	100.0%	100.0%	100.0%	2.97
	CH04	100.0%	91.67%	10	100.0%	100.0%	100.0%	92.31%	90.00%	96.97%	2.01	100.0%	87.50%	100.0%	95.00%	1.74
	CH05	100.0%	91.67%	9	100.0%	100.0%	83.33%	87.50%	90.00%	91.67%	1.96	100.0%	100.0%	100.0%	100.0%	2.11
	Mean	**100.0%**	**95.00%**	**8.60**	**100.0%**	**95.00%**	**96.67%**	**90.96%**	**96.00%**	**94.23%**	2.32	**98.46%**	**94.17%**	**92.00%**	**95.75%**	**2.27**
	SD	0.00%	4.56%	2.07	0.00%	11.18%	7.45%	10.39%	5.48%	7.39%	0.51	3.44%	8.12%	17.89%	4.48%	0.53
MS	CP01	87.50%	79.17%	10	100.0%	100.0%	77.78%	66.67%	90.00%	79.54%	1.52	92.31%	†44.44%	100.0%	76.92%	2.21
	CP02	91.67%	87.50%	6	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	2.09	91.67%	85.71%	100.0%	92.00%	2.53
	CP03	<70%	75.00%	15	†25.00%	85.71%	†57.14%	†-	100.0%	61.11%	1.34	†100.0%	†62.50%	100.0%	85.00%	1.51
	CP04	79.17%	95.83%	13	100.0%	100.0%	100.0%	92.86%	87.50%	96.97%	2.04	100.0%	100.0%	100.0%	100.0%	2.12
	CP05	<70%	<70%	-	-	-	-	-	-	-	-	-	-	-	-	-
	CP06	83.33%	<70%	15	87.50%	72.73%	77.78%	75.00%	87.50%	77.27%	1.72	77.78%	85.71%	85.71%	82.61%	1.59
	CP07	83.33%	91.67%	7	72.73%	†40.00%	83.33%	†50.00%	60.00%	63.33%	2.13	†100.0%	88.89%	100.0%	95.65%	2.69
	CP08	83.33%	70.83%	6	100.0%	100.0%	88.89%	100.0%	90.00%	97.20%	1.66	70.00%	81.82%	75.00%	76.00%	2.76
	CP09	75.00%	95.83%	10	100.0%	†70.00%	100.0%	†62.50%	50.00%	82.35%	1.76	†83.33%	†92.86%	†75.00%	86.67%	-
	CP10	91.67%	75.00%	13	100.0%	†66.67%	75.00%	†62.50%	100.0%	76.00%	1.61	†63.64%	†71.43%	†75.00%	68.18%	-
	CP11	<70%	<70%	-	-	-	-	-	-	-	-	-	-	-	-	-
	CP12	<70%	70.83%	9	†66.67%	72.73%	63.64%	†53.33%	90.00%	62.79%	1.47	†33.33%	†75.00%	†83.33%	72.00%	-
	CP13	83.33%	<70%	8	†71.43%	88.89%	100.0%	80.00%	90.00%	83.78%	2.27	91.67%	100.0%	100.0%	96.30%	2.06
	CP14	87.50%	87.50%	10	75.00%	†75.00%	57.14%	†46.15%	80.00%	62.50%	2.17	†72.73%	100.0%	100.0%	87.50%	2.01
	CP15	91.67%	75.00%	6	†42.86%	70.00%	75.00%	†64.29%	90.00%	64.10%	1.92	†85.71%	†66.67%	†75.00%	75.00%	-
	CP16	<70%	<70%	-	-	-	-	-	-	-	-	-	-	-	-	-
	Mean	**80.45%**	**79.81%**	**9.85**	**80.09%**	**80.13%**	**81.21%**	**71.11%**	**85.77%**	**77.46%**	1.82	**81.71%**	**81.16%**	**83.93%**	**84.14%**	**2.16**
	SD	13.65%	10.60%	3.29	24.48%	17.84%	15.93%	18.67%	14.94%	14.24%	0.30	18.78%	16.69%	11.77%	10.08%	0.44

MS patients CP05, CP11 and CP16 were removed from the assessment because they did not achieved a minimum accuracy of 70% in the calibration sessions.
N_s: number of sequences, acc.: accuracy, OCM: output characters per minute. SD: standard deviation.
[1] Averaged accuracy and OCM include only the tasks in which threshold was enabled (i.e., it does not include the passive reading task, T1.5).
[2] Averaged OCM do not include those tasks that users could not accomplish, since total duration would be unknown.
† Task not completed.

Figure 5.7: Questionnaire results for the assistive web browser application. Boxplots of HS (green) and MS patients (orange) are shown, where mean values are indicated as black dots. For visualization purposes, negative questions are detailed first, followed by positive ones.

5.4 Control of smartphone-based social networks

The experience acquired developing the assistive web browser led to the designing of an asynchronous and assistive P300-based BCI mobile application for controlling smartphone-based social networks, such as Twitter and Telegram. As shown in figure 5.8, the system is mainly composed by three entities: (1) the user, which involves the recording of the EEG signals; (2) the laptop, which receives and processes them, decodes user's intentions and translates them into application commands; and (3) the smartphone, which receives the commands via Bluetooth and offers feedback to the user (Martínez-Cagigal et al., 2019a). Details concerning signal acquisition, pre-processing, and feature extraction are described in sections 3.3, 4.1, and 4.3, respectively. An SW-based LDA classifier was used, according to sections 4.4 and 4.5; as well as a thresholding approach to reach an asynchronous control of the system, previously detailed in section 4.6.1.

As can be observed in figure 5.8, a nested selection matrix approach was em-

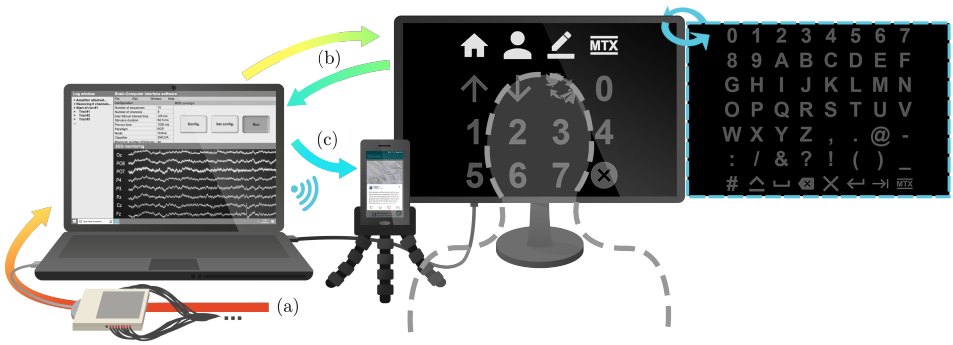


Figure 5.8: Setup of the assistive social networks mobile application from the point of view of the user: (a) EEG signals reach the laptop, which pre-processes them and generates the stimuli; (b) the panoramic screen displays the flashings according to the currently selected matrix; finally, (c) the system determines the selected command and sends it to the smartphone via Bluetooth, which offers feedback to the user.

ployed. Following the same rationale as for the web browser application, it is composed by a small navigation matrix (4×4), intended to operate the application; and a big keyboard matrix (8×7), intended to write texts. As shown, the navigation matrix comprises the numbers from 0 to 7, six functionality buttons (home, profile, write, cancel, scroll up and down), and two buttons for toggling between Twitter/Telegram and navigation/keyboard matrices. Functionalities that could not been adapted using the buttons were tagged in order to be accessed by typing their labels. Figure 5.9 displays six snapshots of the mobile application, including Twitter's timeline and *tweets* view and writing; as well as Telegram's chats, conversation and contacts lists. Signal processing and stimuli displaying were coded with C++ under the BCI2000 platform (Schalk et al., 2004), whereas the mobile application was coded in Java using the Twitter and Telegram open source APIs.

The assistive social networks mobile application was assessed by 18 motor-disabled patients and 5 HS, as detailed in section 3.1. Each user carried out three different sessions: two calibration sessions (Cal-I and Cal-II), and one evaluation session (Eval-I). Calibration sessions followed the exact same experimental procedure as in the web browser assessment (detailed in the previous section 5.3). By contrast, Eval-I was composed by a total of 6 tasks that required a comprehensive control of the BCI application. Table 5.5 details all of them, including the optimal number of selections, the matrices required to accomplish them and the average duration. Note that the duration of each task varied among subjects due to their different optimal number of sequences. Reached accuracies and optimal

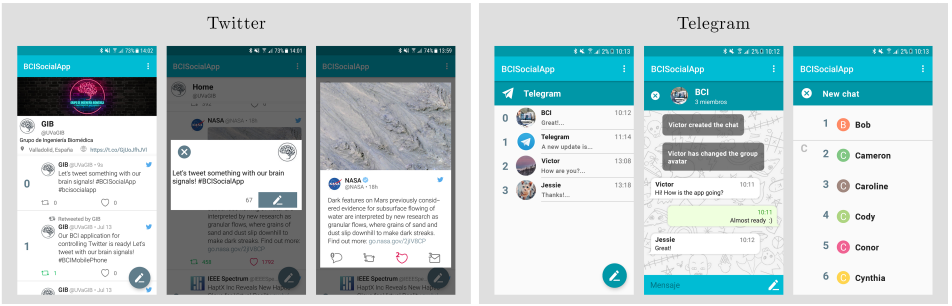


Figure 5.9: Snapshots of the assistive social networks mobile application. From left to right: Twitter's profile timeline, dialog for writing *tweets* and *tweet* view; and Telegram's conversation list, group chat and contacts list.

Table 5.5: Evaluation tasks for the assistive social networks mobile application.

	Task	AT	Description	OS	RM	MDP	
						HS	MS
Eval-I	Toggling	Yes	Scroll up and down Twitter's timeline and toggle to Telegram	3	N	1:10	1:46
	Retweeting	Yes	Scroll down Twitter's timeline, select a *tweet* and *retweet* it	4	N	1:50	3:13
	Writing a *tweet*	Yes	Open the form to write a *tweet* and spell "hello"	7	N,K	3:54	6:01
	Replying	Yes	Access the user profile and reply the last *tweet* with "great!"	11	N,K	5:53	8:13
	Creating a chat	Yes	Select a Telegram's contact, create a new chat and spell "how are you?"	11	N,K	6:15	7:31
	Chatting	Yes	Select a Telegram's chat and reply with "fine, and you?"	12	N,K	7:31	7:49

AT: asynchronous threshold enabled?, OS: optimal selections, RM: required matrices (N: navigation matrix, K: keyboard matrix), AD: average duration in minutes:seconds, HS: healthy subjects, MDP: motor disabled patients.

number of sequences for each session and subject are detailed in table 5.6. It is noteworthy that 4 patients (DP01, DP07, DP11 and DP13) were discarded from the assessment because they did not obtain a minimum accuracy of 70% in the calibration sessions. At the end of Eval-I, a 7-point Likert questionnaire was given to the subjects in order to collect their impressions and suggestions.

Averaged qualitative results of the questionnaires are shown in figure 5.10. As shown, subjects were generally satisfied with the application. Exceptions were the same as in the web browser assessment, including the time that was required to navigate through the application and the toughness of the sessions. By contrast, suggestions were focused on getting rid of the conductive gel and demanding more speed.

Table 5.6: Accuracy results of the assessment sessions for the assistive social networks mobile application.

Subjects		Cal-I	Cal-II	N_s	Eval-I						Average	
					T1	T2	T3	T4	T5	T6	Acc.	OCM[1]
HS	DH01	100%	91.7%	11	100%	100%	100%	90.9%	90.9%	91.7%	93.8%	1.38
	DH02	100%	97.2%	6	100%	100%	85.7%	100%	100%	100%	97.9%	2.37
	DH03	95.8%	95.8%	13	100%	83.3%	85.7%	100%	100%	92.3%	94.2%	1.10
	DH04	100%	95.8%	7	100%	66.7%	100%	81.8%	100%	73.3%	85.2%	1.95
	DH05	87.5%	91.7%	5	100%	100%	100%	90.9%	100%	100%	98%	2.90
	DH06	91.7%	91.7%	8	100%	100%	71.4%	100%	100%	66.7%	86.7%	1.57
	DH07	95.8%	100%	8	100%	60%	57.1%	81.8%	91.7%	81.8%	79.6%	1.84
	DH08	77.8%	91.7%	4	100%	100%	100%	100%	90.9%	91.7%	95.8%	3.55
	DH09	100%	100%	8	100%	100%	100%	91.7%	100%	100%	98%	1.90
	DH10	100%	95.8%	7	100%	80%	100%	100%	90.9%	91.7%	93.9%	2.06
	Mean	**94.9%**	**95.1%**	**7.7**	**100%**	**89%**	**90%**	**93.7%**	**96.4%**	**88.9%**	**92.3%**	**2.06**
	SD	7.4%	3.4%	2.7	0%	14.8%	14.4%	7.1%	4.4%	10.9%	6%	0.73
MD	DP01	<70%	<70%	-	-	-	-	-	-	-	-	-
	DP02	41.7%	83.3%	10	†66.7%	60%	66.7%	63.6%	63.6%	†100%	65.2%	1.58
	DP03	50%	50%	14	†100%	†57.1%	-	-	-	-	72.7%	1.41
	DP04	95.8%	95.8%	9	100%	100%	100%	100%	77.8%	†100%	95.1%	1.51
	DP05	95.8%	70.8%	7	100%	100%	85.7%	90.9%	†100%	100%	95.6%	2.11
	DP06	83.3%	77.8%	7	100%	85.7%	100%	100%	100%	84.6%	94.3%	2.18
	DP07	<70%	<70%	-	-	-	-	-	-	-	-	-
	DP08	87.5%	68.2%	10	100%	100%	85.7%	58.3%	†40%	†71.4%	71.1%	1.50
	DP09	100%	72.2%	13	100%	100%	100%	81.8%	†50%	0%	84.4%	1.15
	DP10	79.2%	79.2%	13	†66.7%	40%	75%	†63.6%	0%	0%	63%	1.20
	DP11	<70%	<70%	-	-	-	-	-	-	-	-	-
	DP12	83.3%	87.5%	12	†66.7%	100%	75%	81.8%	†60%	0%	76.3%	1.15
	DP13	<70%	<70%	-	-	-	-	-	-	-	-	-
	DP14	66.7%	58.3%	9	†66.7%	100%	85.7%	58.3%	†66.7%	†60%	68.8%	1.62
	DP15	83.3%	87.5%	13	100%	66.7%	87.5%	72.7%	100%	100%	88.2%	1.12
	DP16	95.8%	87.5%	14	†66.7%	100%	75%	90.9%	91.7%	100%	89.8%	1.02
	DP17	50%	33.3%	15	100%	83.3%	66.7%	45.5%	-	-	65.5%	1.01
	DP18	95.8%	91.7%	7	100%	100%	100%	100%	100%	92.3%	98%	2.02
	Mean	**79.2%**	**74.5%**	**10.9**	**88.1%**	**85.2%**	**84.8%**	**77.5%**	**70.8%**	**67.4%**	**80.6%**	**1.47**
	SD	18.8%	16.9%	2.7	16%	19.9%	12%	17.7%	29.5%	40.8%	12.4%	0.40

Patients DP01, DP07, DP11 and DP13 were removed from the assessment because they did not achieved a minimum accuracy of 70% in the calibration sessions.
N_s: number of sequences, Acc.: accuracy, OCM: output characters per minute. SD: standard deviation.
[1] Averaged OCM do not include those tasks that users could not accomplish, since total duration would be unknown.
† Task not completed.

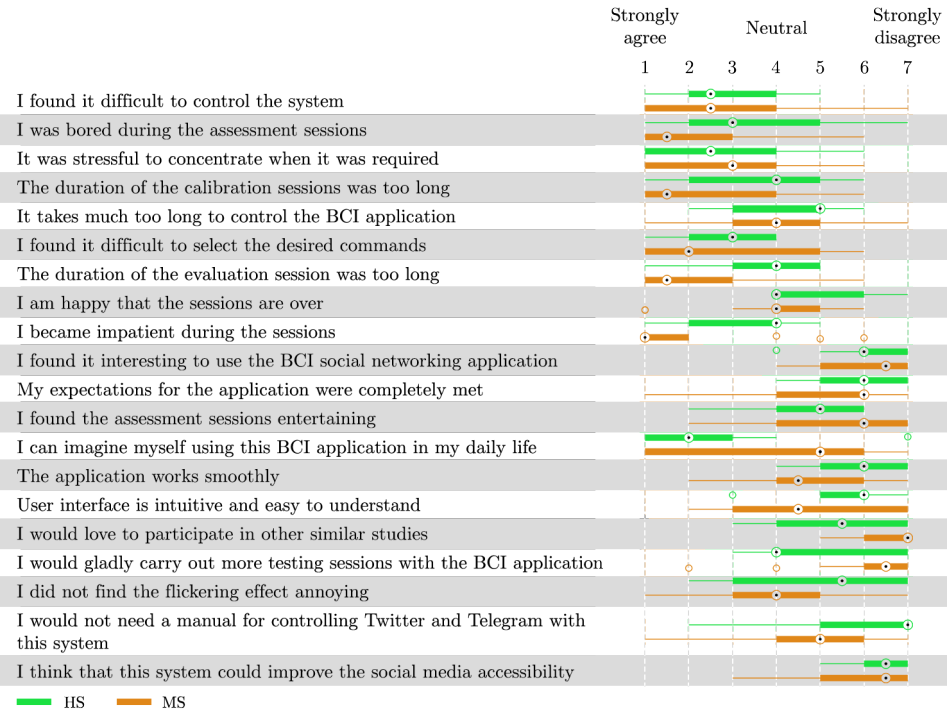


Figure 5.10: Questionnaire results for the assistive social networks mobile application. Boxplots of HS (green) and motor-disabled patients (orange) are shown, where mean values are indicated as black dots. For visualization purposes, negative questions are detailed first, followed by positive ones.

Chapter 6

Discussion

In this Doctoral Thesis, several aspects of practical P300-based BCIs have been addressed. Firstly, BCI asynchrony was explored by proposing two different approaches. The thresholding approach was implemented and tested in a real application with motor-disabled subjects, assuring its efficacy in a practical environment. Then, an offline MSE approach was able to successfully characterize asynchronous states, finding that control signals are significantly more complex and irregular than non-control ones. Secondly, a novel multi-objective meta-heuristic was especially designed to select a customized set of channels for each user. The viability of the proposed method was analyzed in terms of computational cost and convergence, as well as by comparing the results with those yielded by other existing single- and multi-objective algorithms. Thirdly, a practical asynchronous BCI web browser was developed and tested with HS and MS subjects. Testing accuracies of HS and MS patients reached 95.75% and 84.14%, respectively; exhibiting a suitable control of the system. Lastly, a mobile BCI application to access social networks was developed and assessed by HS and motor-disabled subjects. Results showed that users could successfully control the BCI, reaching averaged online accuracies of 92.3% and 80.6%, respectively. We consider that these applications contributed to bridge the accessibility gap of Internet and mobile networks, making an impact in the quality of life of the motor-disabled by improving their personal autonomy. In this chapter, the aforementioned findings are discussed according to the flow of hypotheses and results of previous chapters. At the end, the main limitations of this thesis are presented.

6.1 Assuring the asynchrony of practical systems

Assuring the asynchrony of the BCI is essential to provide a comprehensive self-paced control of the system in a real environment. The thresholding approach has demonstrated its viability in a practical setup, reaching averaged training accuracies above 96% and 84% for HS and motor-disabled patients, respectively; and avoiding a large number of incorrect selections in online sessions (Martínez-Cagigal et al., 2017, 2019a). This implies that the discrimination of non-/control state signals based on SW-LDA scores is feasible. Furthermore, owing to the fact that the method is embedded into the classification stage, it does not increase the overall complexity of the signal processing pipeline. However, its greatest advantage it is also its main defect. It is well known that the EEG presents a high inter-session variability, which causes the need to recalibrate the ERP classifier frequently. The dependence on the classifier forces to also update the thresholding approach due to changes in classifier's scores, which implies recording additional non-control signals. Therefore, its performance is linked to the pattern recognition stage, compromising the overall system as the number of sessions increase.

The aforementioned limitation of the thresholding method is overcome by the MSE approach. It has been demonstrated that SampEn-based MSE can characterize non-/control state signals, allowing a reliable discrimination between them, which is also independent of the ERP classifier (Martínez-Cagigal et al., 2019b). As shown, both states behave similarly to dynamical changes in different MSE scales, but control signals present a steeper slope. Thus, non-control signals are more regular in most scales, which implies that control signals are, from the point of view of the information theory, more complex (Costa et al., 2002). It is noteworthy that non-control signals become more unstable as the scale increases, whereas control ones are generally more defined, showing smaller values of standard deviation. For most subjects, these variations yielded significant differences (p–values < 0.05) in almost all channels, favoring the discrimination between both states. Topographical results also showed an easier discrimination over the PFC, commonly associated with planning complex cognitive behaviors, personality expression, decision making and selective attention (Miller et al., 2002). This is consistent with the RCP task, which implies a constant attention to identify target stimuli among other background stimuli. In fact, *oddball* tasks were demonstrated to produce hemodynamic changes in the dorsolateral PFC, associated with response strategies (i.e., mapping of the stimuli to responses) (Huettel and McCarthy, 2004). Another recent study suggested that the elicitation of P300 waves are linked to complex

processes such as memory, attention and decision making; somewhat related with the PFC (Bojorges-Valdez and Yanez-Suarez, 2018). Although less intense, differences over V1 are also shown. Neurons of V1 fire action potentials whenever visual stimuli appears in the receptive field and thus, a higher amount of neurons are expected to be activated in the control state (Standring, 2016). Since the feature extraction process used raw signals including target and non-target stimuli, neural activation of V1 is expected to spread electrical activity across the frequency spectrum. Entropy measures are thus able to follow that activation, increasing the irregularity of control state signals.

The ability of SampEn to discriminate between asynchronous states is reinforced in the validation sessions, reaching an average accuracy of 94.4% in HS. Even though the asynchronous detection is feasible for all scales, maximum differences were found in $\tau = 2$, equivalent to reducing the sampling rate of the EEG signal by half before applying the SampEn algorithm. In practice, it would not be required to compute all MSE scales, but only to implement a decimation block, followed by a SampEn feature extraction process. As shown, the total computational duration for 15 sequences was 196.8 ms. Since most P300-based BCIs use pauses of more than 1 second between characters (Farwell and Donchin, 1988; Schalk et al., 2004; Treder et al., 2011), the computational cost of the proposed algorithm assures its viability in a practical online BCI assessment. Discrimination accuracies also depicted an increasing tendency as the number of sequences increased, suggesting that the proposed method is dependent on the length of the raw signals. Particularly, all subjects except AH04 and AH08 reached more than 90% accuracy using 9 sequences; and even AH06, AH07 and AH09 reached it using only 3 sequences. Although this tendency is clear for all subjects, some users presented a sequential increase (e.g., AH01, AH05, AH08, AH09), while others reached a standstill (e.g., AH03, AH10). This behavior reinforces the need to perform individual calibrations and optimize BCI applications to each subject separately.

Table 6.1 summarizes a comparison between previous asynchronous P300-based BCI applications, emphasizing their asynchronous strategies. As shown, most of the studies followed a thresholding approach, either LDA (Aloise et al., 2011; Breitwieser et al., 2016; Martínez-Cagigal et al., 2017, 2019a; Pinegger et al., 2015; Tang et al., 2018) or SVM based (Li et al., 2013; Zhang et al., 2008). Despite their suitable performances, the dependence on the ERP classifier makes them impractical in the long term due to the inherent inter-session variability of the EEG. As aforementioned, threshold values should be updated each time the classifier

Table 6.1: Comparison between asynchronous P300-based BCI applications.

Study	Signal	Paradigm	Asynchronous management	Subjects
Zhang et al. (2008)	P300	Cells	ROC thresholding using SVM scores	4HS
Panicker et al. (2010)	P300, SSVEP	Hybrid	Detection of SSVEPs using relative peak amplitude in PSD	10HS
Aloise et al. (2011)	P300	RCP	ROC thresholding using LDA scores	11HS
Li et al. (2013)	P300, SSVEP	Hybrid	ROC thresholding using SVM scores (P300) and relative powers (SSVEP)	8HS
Pinegger et al. (2015)	P300	RCP	Thresholding using LDA scores and sum of spectral components	10HS
Breitwieser et al. (2016)	P300, SSSEP	Hybrid	Thresholding using multi-class LDA	14HS
Martínez-Cagigal et al. (2017)	P300	RCP	ROC thresholding using LDA scores	5HS, 16MS
He et al. (2017)	P300	RCP	Combination of two different SVM	8HS
Yu et al. (2017)	P300, SMR	MI, RCP	SMR activate the RCP	11HS, 8HS
Aref and Huggins (2012)	P300	RCP	Certainty algorithm: t-test over LDA scores	11HS, 19CP
Ma and Qiu (2018)	P300	RCP	ROC thresholding using relative powers	4HS
Aydin et al. (2018)	P300	HOS	ROC thresholding using classifier labels	10HS
Tang et al. (2018)	P300	RCP	ROC thresholding using LDA scores	4HS
Santamaría-Vázquez et al. (2019)	P300, SSVEP	RCP	Oddball steady response detection	15HS
Martínez-Cagigal et al. (2019a)	P300	RCP	ROC thresholding using LDA scores	18HS, 10MDS
Martínez-Cagigal et al. (2019b)	P300	RCP	SampEn features and LDA classification	10HS

SSVEP: steady-state visual evoked potentials, SSSEP: somatosensory evoked potentials, SMR: sensory-motor rhythms, MI: motor imagery, RCP: row–col paradigm, ROC: receiver operating characteristic, SVM: support vector machines, PSD: power spectral density, LDA: linear discriminant analysis, SampEn: sample entropy, CS: control subjects, MS: multiple sclerosis, CP: cerebral palsy, MDS: motor-disabled subjects.

weights are modified, requiring additional EEG recordings of non-control state observations. Other approaches added complementary spectral features (Ma and Qiu, 2018; Pinegger et al., 2015) or employed hybrid paradigms (Li et al., 2013; Panicker et al., 2010; Yu et al., 2017) to implement filter methods independent of the ERP classifier. Yu et al. (2017) used SMR to activate RCP stimuli via MI, including a selection command to stop them. However, SMR-based control is considered unreliable and the extra command selection entails an increase of the required time to handle the asynchrony, making the system even more demanding. Li et al. (2013) and Panicker et al. (2010), by contrast, superimposed the RCP with a flickering visual effect. This strategy was intended to generate SSVEPs whenever users were paying attention to the stimuli, further detected using relative powers. An utterly different approach was followed by Pinegger et al. (2015) and Ma and Qiu (2018), who hypothesized that the RCP also generates residual

SSVEP components provided the stimuli are displayed at a constant rate. These SSEVPs were identified in the frequency spectrum, providing supplementary features to the LDA scores. Recently, Santamaría-Vázquez et al. (2019) extended this approach by proposing a novel filter method based on canonical correlation analysis that does not require the recording of non-control trials. Unfortunately, most of these studies were intended to develop novel assistive applications, reporting accuracies that depict the overall performance to predict correct characters, but ignoring non-control ones. For that reason, a quantitative statistical comparison cannot be performed among them.

To sum up, one wrapper (thresholding) and one filter (MSE) asynchrony methods have been proposed. The thresholding approach has demonstrated its viability with motor-disabled subjects, yet it is strongly dependent on inter-session variability (Martínez-Cagigal et al., 2017, 2019a). Although this fact does not invalidate thresholding-based approaches straightaway, it is time consuming, affects the usability of the BCI and could be frustrating for users. Hence, strategies that do not depend on the classifier scores, such as residual SSVEP detection or the proposed entropy-based algorithm, should be preferred in the long-term. Although none of them has been evaluated with a motor-disabled population, results obtained by HS suggest the generalization is feasible. Among these methods, the SampEn-based has demonstrated to be able to characterize non-/control signals and monitor their dynamical changes to provide an asynchronous management independent of the inter-session variability (Martínez-Cagigal et al., 2019b). It is also believed that performance could be improved by integrating other features from complementary methods, such as spectral-based algorithms; however, this hypothesis has not been explored yet.

6.2 A novel nature-inspired algorithm to select relevant channels

Notwithstanding their still scarce application in BCIs, it has been demonstrated that single- and multi-objective meta-heuristics are able to select optimal BCI channel sets according to the features of each user, independent of the P300-based paradigm (Martínez-Cagigal et al., 2020). In further BCI sessions, the supervisor should use the selected channels to avoid over-fitting, while reducing the preparation time and the overall cost of the assessment.

In the light of the results, single-objective meta-heuristics such as GA, BDE

and BPSO have succeeded in finding proper channel sets. In training phase, GA converged the fastest, followed by BDE and finally BPSO. Moreover, the overall complexity of the algorithms also follows this behavior, being GA the less computationally intensive, followed by BDE and BPSO. However, the quality of their optimized solutions $F(\boldsymbol{s})$ were analogous, which explains the similarity of their performances in testing phase. This behavior is consistent for all databases, reinforcing the generalization of the results across different paradigms. It could be thus considered that GA is more fitted to the BCI channel selection problem, likely due to its inherently discrete inner strategies. On average, GA, BDE and BPSO reached 91.47% accuracy using 13.38 channels. The increase in accuracy compared to the entire channel set (i.e., ALL) and the standard Krusienski's P300 8-channel set (i.e, KRU) was statistically significant for both (p-value < 0.05). Therefore, it can be assured that customized channel sets are able to significantly improve the performance, making channel selection a beneficial stage for BCI system. It should be noted, however, that single-objective sets employed more channels than KRU. In this regard, channel densities of the solutions may be easily modified by adjusting the importance of the aggregated objectives: minimization of the amount of channels, $f_1(\boldsymbol{s})$; and maximization of the overall performance, $f_2(\boldsymbol{s})$ (equation 4.1).

Despite the possibility to tune the total number of channels to be selected, single-objective meta-heuristics always return a single solution. Multi-objective algorithms, by contrast, return a set of optimal solutions, each one with different number of channels. The supervisor is then allowed to select one of them in function of the desired channel density. This fact poses a major advantage over traditional single-objective methods that should be taken into account if the supervisor has enough time for the computation. Among the existing algorithms, NSGA-II and SPEA2 reached similar Pareto-fronts, outperforming BMOPSO. Nonetheless, some of their inner strategies are suboptimal in the BCI channel selection problem, hindering the convergence of the algorithms. In this framework, solutions must be dichotomous, making continuous-based operations futile or counter-productive, such as crowding distances, distance soughts or transfer functions. Even repository control approaches are worthless, since its maximum size is already limited by the number of channels. Ultimately, all meta-heuristics must have been adapted to the problem to a greater or lesser extent. DFGA, by contrast, have been especially designed for the BCI channel selection framework, and so its Pareto-front is comparable to NSGA-II and SPEA2 while it converges faster. Channel ranks of BMOPSO also exposed a lack of convergence to a global optimum, which was

afterward reflected in testing results. Accuracies indicated that KRU was outperformed using only 4 channels by DFGA, NSGA-II and SPEA2. Furthermore, results showed that there was a point where performance came to a standstill, making the use of more than 15–20 channels counter-productive.

From the channel ranks results it was observed that optimal channel sets differ for each subject, confirming that a customized set benefits the subsequent performance. In spite of the usefulness of KRU as a general rule-of-thumb solution (Krusienski et al., 2008), results did not consider that combination optimal for any case. In fact, KRU was outperformed by both single and multi-objective meta-heuristics, even using only 4 channels in case of DFGA, NSGA-II and SPEA2. This fact should not be surprising, since classifiers are always optimized for each subject due to the inherent variability of the EEG signals (Wolpaw and Wolpaw, 2012a). In fact, the inter-subject variability of optimal channel sets have been repeatedly observed in the literature, suggesting that the concept of using a general channel set for an entire population is suboptimal and should be avoided to the extent possible (Chaurasiya et al., 2017; Gonzalez et al., 2013; Jin et al., 2010; Kee et al., 2015). Since the selection of channels is somewhat related with the identification of P300 waves, a possible explanation of this phenomenon could be related with inter-subject differences between these potentials. It is documented that the amplitude, latency, and even the topographic distribution of P300 waves vary depending on: (i) individual differences (e.g., age, intelligence, personality, absolute pitch), (ii) psychological (e.g., attention, motivation) and (iii) pharmacological aspects (e.g., consumption of alcohol and other drugs), (iv) the paradigm (e.g., task complexity, target probability), (v) psychiatric disorders (e.g., schizophrenia, autism, obsessive-compulsive disorder, psychopathy, dementia, etc.), and (vi) external factors (e.g., variations in cap positions) (Picton, 1992). In fact, the latter is present in almost all recordings, where EEG caps do not correctly fit some users, making some electrodes wobbly and producing noise. Taking into account all these facts, it is not surprising that most relevant channels for classification are not the same among different users, requiring a subject-optimized procedure to maximize the BCI performance. Moreover, channel ranks showed that meta-heuristics had a slight tendency to select electrodes over the occipital lobe. From a biological point of view, this tendency is sound. The EEG response of these visual *oddball* task is modeled as an ERP, composed by several components that are taken into account when extracting and classifying the features. Due to the ability of V1 to process information about visual stimuli, static and moving objects, it is thus expected that occipital electrodes contain relevant discriminative information about target

and non-target responses (Standring, 2016).

The main drawback of these meta-heuristics refers to the need of fixing hyperparameters whose values usually depends on the context of the optimization problem. Convergence issues may arise if these values are poorly chosen. Thus, the quality of a meta-heuristic should be assessed taking into account the amount of required hyperparameters: the less parameters, the more probability to assure a generalization of the results. In this sense, GA, DFGA, NSGA-II and SPEA2 only require mutation and crossover probabilities to be fixed, which are well studied in the literature and generally take values of $1/N$ and 0.90–0.95, respectively (Deb, 2005; Deb et al., 2002; Yang, 2014). A similar approach is followed by BDE, whose extra parameters are intended to apply a mutation procedure. However, BPSO and BMOPSO add the maximum velocity, as well as the personal and global confidences, which are problem-dependent. Although we followed the recommendations of the literature, results expose a clear difficulty of BPSO and BMOPSO to converge to an optimal solution. This fact suggests that their performances could have been improved by means of an adaptive approach or a hyperparameter optimization.

An analysis of the computational cost of these algorithms was also performed, although its interpretation is tricky, forcing to consider several aspects at the same time. The number of evaluations per generation varied depending on the algorithm and thus, the meta-heuristics can only be compared in terms of the duration of a single generation. In order to assure a fair comparison, all the algorithms performed a total of 4000 evaluations, which implies that the number of generations of each algorithm varied. However, the total time of execution varies in practice according to the required number of generations to achieve a suitable convergence, the search depth and the programming approach. Since an intense search would inevitably generate repeated solutions, the use of hash maps is mandatory to match the computed solutions with their fitnesses and avoid unnecessary evaluations. In terms of asymptotic operations, all of the multi-objective meta-heuristics showed a $O(N_o m^2)$ behavior (Curry and Dagli, 2014; Martínez-Cagigal et al., 2020; Tripathi et al., 2007), whereas the number of evaluations per generation made NSGA-II the less time-consuming, followed by DFGA, SPEA2 and, finally, BMOPSO. The scarce computational cost of DFGA and NSGA-II, together with their performances in testing phase, makes them excellent multi-objective approaches to this problem. For the single-objective ones, GA and BDE stood out owing to their convergence abilities and their low computation costs. In any case, the overall duration of these algorithms restricts their application to the

Table 6.2: Comparison between channel selection meta-heuristics in P300-based BCIs.

Study	Database	N_s	Method	Accuracy	N_c
Kee et al. (2015)	Comp.	15	GA	93.60%	22.3
			NSGA-II	94.90%	25.7
Arican and Polat (2020)	Comp.	15	BPSO	89.90%	8
Perseh and Sharafat (2012)	Comp.	15	BPSO	85.00%	31
Gonzalez et al. (2013)	Comp.	5	BPSO	67.50%	33.5
Martínez-Cagigal and Hornero (2017b)	Comp.	15	BPSO	92.00%	17
			BA	96.00%	19
			ABC	93.00%	24
			BAS	92.00%	19
			FA	94.00%	22
Chaurasiya et al. (2017)	Custom, 9HS	15	MOBDE	92.80%	26.1
Jin et al. (2010)	Custom, 11HS	15	BPSO	71.09%	7.63
Martínez-Cagigal et al. (2020)	Comp.	15	GA	92.00%	14
			BDE	92.00%	14.5
			BPSO	92.00%	14
			DFGA*	94.50%	20
			NSGA-II*	94.50%	20
			SPEA2*	94.00%	16
			BMOPSO*	92.50%	20
	Center	10	GA	97.40%	12.4
			BDE	97.90%	12.5
			BPSO	96.80%	12.5
			DFGA*	97.88%	7
			NSGA-II*	98.46%	8
			SPEA2*	97.72%	9
			BMOPSO*	97.82%	16
	RSVP	10	GA	84.60%	13.4
			BDE	85.50%	13.4
			BPSO	85.00%	13.7
			DFGA*	85.73%	14
			NSGA-II*	86.05%	7
			SPEA2*	85.73%	8
			BMOPSO*	84.80%	18

N_s: number of sequences, N_c: averaged number of channels, Comp.: III BCI Competition 2005 (dataset II), HS: healthy subjects, GA: genetic algorithm, NSGA-II: non-sorting genetic algorithm 2, BPSO: binary particle swarm optimization, BA: bees algorithm, ABC: artificial bee colony, BAS: binary ant system, FA: firefly algorithm, MOBDE: multi-objective binary differential evolution, BDE: binary differential evolution, DFGA: dual-front sorting algorithm, SPEA2: strength pareto evolutionary algorithm 2, BMOPSO: binary multi-objective particle swarm optimization.
* The selected solution was the one that maximized the accuracy in the range $N_c \in [5, 20]$.

calibration session, where the classifier's weights are optimized for each subject.

According to the outcomes, the utility of meta-heuristics (GA, BDE, DFGA, NSGA-II, SPEA2) to find an optimal combination of channels in P300-based BCIs have been proven. In fact, the reached accuracies of our work are similar or even higher than those reported previously. Table 6.2 depicts a comparison between previous channel selection meta-heuristics in P300-based BCIs. The most straightforward comparison can be made by considering accuracies from those that employed the 'BCI Competition III: dataset II'. Kee et al. (2015) reached 93.60% (22.3 ch., GA) and 94.90% (25.7 ch., NSGA-II). Arican and Polat (2020) reached 89.9% (8 ch., BPSO). In our preliminary work, we obtained 92.00% (17 ch., BPSO), 96.00% (19 ch., BA), 93.00% (24 ch., ABC), 92.00% (19 ch., BAS) and 94.00%

(22 ch., FA) (Martínez-Cagigal and Hornero, 2017b). All of these studies used 15 sequences. Perseh and Sharafat (2012) and Gonzalez et al. (2013) obtained 85.00% (31 ch., BPSO) and 67.50% (33.5 ch., BPSO), respectively, using only 5 sequences. As can be seen, the comparison is difficult since each study reported solutions with different number of channels or sequences. In our study, we reached 92.00% (14 ch., GA, BPSO; 14.5 ch., BPSO), and NSGA-II, SPEA2 and DFGA achieved 90.00% with 7 channels, which increased until a maximum of 97.00% using 23 channels and 15 sequences (Martínez-Cagigal et al., 2020). There were also studies that tested their proposals with custom datasets, such as Chaurasiya et al. (2017) (9HS, 15 sequences) and Jin et al. (2010) (11HS, 15 sequences), obtaining 92.80% (26.1 ch., MOBDE) and 71.09% (7.63 ch., BPSO), respectively. Besides the 'BCI Competition III: dataset II' (2HS, 15 sequences), our study also reported results with two additional databases: 'Center Speller (008-2015)' (13HS, 10 sequences) and 'RSVP Speller (010-2015)' (12HS, 10 sequences). To the best of our knowledge, there are no previous studies that have tested any meta-heuristic with any paradigm apart from RCP and thus, a direct comparison cannot be made. In terms of yielded accuracies, our results for single-objective (averaged, CS: 97.36%, 12.46 ch.; RSVP: 85.03%, 13.5 ch.) and multi-objective (averaged, CS: 98.46%, 8 ch.; RSVP: 85.73%, 8 ch.) algorithms in these databases are similar to those reported in the literature (Acqualagna and Blankertz, 2013; Treder et al., 2011).

To sum up the main findings, it was found that optimal channel sets showed a high inter-subject variability, making the customization for each user essential regardless of the employed paradigm. Moreover, inherently discrete algorithms (DA, BDE, DFGA, NSGA-II, SPEA2) reached higher performances due to the discrete nature of the framework. It is also noteworthy that the combination of deterministic and stochastic approaches seemed to be beneficial for the convergence of the algorithm, as shown by the proposed DFGA.

6.3 A web browsing application for real users

The most immediate BCI application is to somehow improve the quality of life of those whose motor abilities are restricted. Unfortunately, BCI studies often lack of a clinical validation with target users. In order to contribute to the practical BCI literature, we proposed a novel asynchronous web browser to assist MS patients (Martínez-Cagigal et al., 2017). The web browser integrated a Google Chrome extension that interpreted the commands selected by two RCP-based matrices, while using an asynchronous thresholding approach.

In the evaluation sessions, HS and MS patients reached averaged accuracies of 95.75% and 84.14%, which proves the feasibility of the proposed application. As expected, HS obtained higher accuracies with a fewer number of sequences than MS patients, so the surfing speed of HS was higher. It is noteworthy that 3 MS patients (CP05, CP11, CP16) were removed from the evaluation due to low calibration accuracies. This phenomenon is common in assessments with target users because of poor P300 potentials (e.g., attenuated or even null response, variable latencies) (Wolpaw et al., 2002), which again reinforces the importance of validating BCI studies in real environments. It should be also taken into account that some of these MS patients presented neurological damage that caused cognitive disability besides motor limitations. As indicated in section 3.1, this led to different degrees of sustained attention, a critical aspect for triggering P300 responses (Picton, 1992). Furthermore, although HS finished all tasks, not all MS patients were able to finish them, which reflects a poorer control of the system. However, 13 MS patients obtained accuracies greater than 80%, of which 2 did not perform any mistake (Eval-I: CP02; Eval-II: CP04). All HS obtained accuracies greater than 80%, and 3 of them reached a perfect control of the browser (Eval-I: CH01, CH03; Eval-II: CH03, CH05). The comparison between both evaluation sessions also indicates an interesting fact about the thresholding approach. Eval-II (i.e., asynchronous) accuracies are higher than Eval-I (i.e., synchronous) ones for both groups of participants, reaching an average improvement of 6.68% for MS patients. This phenomenon suggests that, on subjects without motor or cognitive capabilities, the introduction of the asynchronous approach does not imply an improvement in terms of performance, but a less demanding control.

It should be noted that a bad optimized threshold can lead to a longer required time to accomplish the tasks due to false negative errors. This fact was present for CP02, who got perfect performances in T1.3 and T2.2 (same task), but the required time to finish T2.2 was longer because 10% of the selections were false negatives. As aforementioned in section 6.1, this problem was caused by the inability of the threshold to follow nonstationary changes of the EEG. Thus, the thresholding approach establish a tradeoff between browsing speed and selection accuracy. Despite this phenomenon, performances were improved when threshold was enabled, allowing users to avoid further mistakes when their P300 are not powerful enough for being considered deliberate selections.

Questionnaire results lay bare the general satisfaction of HS and MS with the web browsing application. They found the system interesting, intuitive and they stated to be willing to participate in further studies. MS patients also indicated

Table 6.3: Comparison between P300-based BCI web browsers.

Study	Signal	Selection strategy	Asynchronous management	Subjects	Acc.
Karim et al. (2006)	SCP	Dichotomous tree	None	1ALS	80.00%
Bensch et al. (2007)	SCP, SMR	Dichotomous tree	None	4ALS 2HS	N.r. N.r.
Mugler et al. (2010)	P300	Node tagging	Pause command	3ALS 10HS	72.00% 93.40%
Sirvent Blasco et al. (2012)	P300	Cursor	Pause command	4HS	93.00%
Yu et al. (2012)	P300, SMR	Cursor	Absence of MI	7HS	93.21%
Martínez-Cagigal et al. (2017)	P300	Node tagging	Thresholding	16MS 5HS	84.14% 95.75%

Acc.: accuracy, SCP: slow cortical potential, SMR: sensorimotor rhythms, MI: motor imagery, ALS: amyotrophic lateral sclerosis, HS: healthy subjects, MS: multiple sclerosis, N.r.: not reported.

that they could imagine themselves using the BCI web browser in their daily life. The main concern of MS patients was the navigation speed. However, browsing speed is directly related with the optimal number of sequences for each user and, in the end, with the classifier's training accuracy. This issue does not appear in HS responses, likely because they committed fewer mistakes and used fewer number of sequences than MS. Participants also pointed out that they were sometimes distracted by adjacent flashings. This issue, known as the 'adjacent-distraction problem', is inherent to the RCP and cause selection errors to fall in cells that belongs to the same row or column as the target command. In our evaluation, the 100% and the 87.75% of the mistakes of HS and MS were of this kind, respectively. Since the probabilities of randomly selecting one of those cells are between 45%–67% (navigation matrix) or 29%–36% (keyboard matrix), it is clear that most errors were due to this problem. Whether a modified RCP, such as the checkerboard paradigm, could have reduced this number is still an open question (Townsend et al., 2010).

In summary, the main strengths of our proposal were: (i) the reliability that guarantees P300 potentials; (ii) the node tagging and double matrices strategies that increases the navigation speed; and (iii) the asynchronous mode. Table 6.3 shows a comparison with previous BCI web browsers, in terms of control signals, selection strategy, asynchronous management and testing accuracies. Besides the use of P300 potentials and node tagging, which makes our proposal faster and more self-sufficient than other SCP/SMR-based BCIs (Bensch et al., 2007; Karim et al., 2006), the main advantage was the asynchronous approach. Previous studies opted to implement a "read mode" command, which paused the application for a fixed time (Mugler et al., 2010; Sirvent Blasco et al., 2012); used the absence of MI to stop involuntary selections (Yu et al., 2012); or directly avoided the implemen-

tation of any asynchronous management (Bensch et al., 2007; Karim et al., 2006). However, a "read mode" command increase the rigidity of the application, whose pause can result too long or too short a time depending on the situation; and a simultaneous SMR-based hybrid control makes the system even more demanding. In our approach, users could experiment a free surfing without worrying about undesired selections from the RCP, due to a continuous attention monitoring through our thresholding method. Although we observed a highly variable performance of MS patients during the sessions, the yielded accuracy (84.14%) is higher than those reported by previous attempts tested with ALS patients (Karim et al., 2006; Mugler et al., 2010). Specifically, significant differences were found between our MS results and the work of Mugler et al. (2010) (p–values < 0.05), which reinforces the contribution of our work considering that the cognitive disabilities that commonly appear in MS are rarely presented on ALS patients. HS results (95.75%) also overcame previous testing with healthy participants, although the difference was not statistically significant (Mugler et al., 2010; Sirvent Blasco et al., 2012; Yu et al., 2012). These results reveal that the inclusion of an asynchronous management significantly favors the applicability of the BCI system in a real environment.

6.4 Toward smartphone-oriented BCIs

Nowadays, smartphones constitute an important aspect of our lives. Their functionalities, that cover from managing finances to reading news, including watching videos, shopping, playing games or searching for information; facilitate many aspects of the everyday life. Particularly, more than the 56% of the time spent with smartphones is dedicated to socializing (Ipsos MORI and Google, 2017) and, unfortunately, the accesibility of these devices is still restricted for those that cannot use accurately their hands and fingers. Therefore, we studied whether the integration of BCI to control an smartphone is feasible (Martínez-Cagigal et al., 2019a).

A total of 4 MDS were discarded from the assessment because they could not achieve a minimum of 70% accuracy in the calibration sessions. Since the diseases were heterogeneous, the rationale behind this lies in indirect issues that caused poor P300 responses. For instance, some users exhibited lack of sustained attention, essential tremors, or nystagmus, among others; which indeed affects the general performance of the system. It is noteworthy that this fact is always present when testing with target users owing to the high inter-subject variability of their symptoms, even for those that present the same disease. It also empathizes

that evaluating with HS and patients are radically distinct experiences. From the experience acquired by testing both the web browser and the social networking applications, it could be suggested that a patient tested BCI cannot be fully generalized to other patients. Hence, the viability of a HS-tested BCI should not be guaranteed for disabled patients under no circumstances.

As indicated in the results of the evaluation session, HS achieved an averaged accuracy of 92.3%, whereas motor-disabled patients reached 80.6%. These results are in accordance with the web browser assessment, which stated that patients had a big handicap when controlling the BCI system, in comparison with HS. Motor-disabled subjects also experimented a lower speed because of their higher number of sequences, and not all of them were able to finish all the tasks. Differences between accuracies, number of sequences and average OCM of HS and MDS were significant (p–values < 0.05). Despite the aforementioned issues, 80.6% is considered sufficient for experimenting an actual control of a BCI (Kübler et al., 2001). Noteworthy, it was observed that selection errors often caused more mistakes thereafter, probably due to despondency. A possible bypass could rely on spelling dictionaries or processing error-related potentials.

Questionnaire results were similar to those obtained with the web browser, showing a general satisfaction with the system. Users did not experiment impatience, boredom, fatigue or stress; and patients could imagine themselves using the application in their daily life. However, they demanded more speed and shorter evaluation sessions. The latter also reveals an issue that should be taken into consideration when designing the tasks, their duration and the structure of the assessment sessions; which could be tiresome for some users due to individual characteristics caused by their diseases. Even so, users were willing to carry out more sessions and participate in similar studies.

Notwithstanding the growing popularity of smartphones, mobile BCIs are scarce in the literature. Furthermore, to the best of our knowledge, there are no previous studies that attempted to provide a high-level control of the device, nor controlling any social network or their functionalities. Table 6.4 shows a comparison among previous mobile BCIs, detailing the control signal, the target operating system, their main functionalities and the final assessment accuracies. Previous studies were focused to dial numbers (Chi et al., 2012; Wang et al., 2011), call contacts (Campbell et al., 2010; Katona et al., 2014; Wang et al., 2011), spell words (Elsawy et al., 2017; Obeidat et al., 2017), visualize the gallery (Elsawy and Eldawlatly, 2015) or play simple games (Wu et al., 2014). Most of them used P300 potentials, whereas some of them used SSVEPs (Chi et al., 2012; Wang et al.,

Table 6.4: Comparison between mobile P300-based BCI applications.

Study	Signal	Target OS	Main functionalities	Subjects	Acc.[1]
Campbell et al. (2010)	P300	iOS	Call contacts	3HS	88.89%
Wang et al. (2011)	SSVEP	Cell	Dial numbers	10HS	95.90%
Chi et al. (2012)	SSVEP	Cell	Dial numbers	2HS	89.00%
Katona et al. (2014)	Conc.	WP	Accept/reject incoming calls	5HS	75.00%
Wu et al. (2014)	Conc.	Android	Play a simple racing game	5HS	N.r.
Elsawy and Eldawlatly (2015)	P300	Android	Open pre-installed apps and visualize the gallery	6HS	79.17% 87.50%
Elsawy et al. (2017)	P300	Android	Spell words	6HS	64.17%
Obeidat et al. (2017)	P300	Android	Spell words	14HS	90.00%
Martínez-Cagigal et al. (2019a)	P300	Android	Full asynchronous control of Twitter and Telegram	18MDS 10HS	80.60% 92.30%

OS: operating system, Acc.: accuracy, SSVEP: steady state visual evoked potential, Conc.: concentration signal (derived from Neurosky EEG caps), Cell: cell phone, WP: Windows phone, HS: healthy subjects, MDS: motor-disabled subjects, N.r.: not reported.
[1] If a study provides several measurements, the highest online accuracy is shown.

2011) and Neurosky concentration (Katona et al., 2014; Wu et al., 2014) control signals. Although the processing of the latter is simple and can be handled by the headset itself, it is a commercial signal that only offers dichotomous decisions, hindering its use for a high-level control of a smartphone. As far as we know, none of them tested their proposals with motor-disabled patients, so their feasibility with target users is compromised. Moreover, the averaged HS accuracy of our approach (92.30%) overcame those reported for P300-based studies (Campbell et al., 2010; Elsawy and Eldawlatly, 2015; Elsawy et al., 2017; Obeidat et al., 2017), and the results of MDS (80.60%) is also higher than those obtained by HS in (Katona et al., 2014) and Elsawy and Eldawlatly (2015). In fact, significant differences in performance were found between the latter and our study (p–values < 0.05). The remaining studies did not provide unfolded accuracies for each user and thus, statistical analysis could not be performed.

As shown, very few studies have attempted to control mobile devices through BCIs, and none of them was focused on providing a high-level control of the device nor social networking services. Our system provided a comprehensive control of Twitter and Telegram, covering all their functionalities and simultaneously reaching high accuracy results. For that reason, we could consider this study as a precursor that demonstrated the viability of smartphone-oriented BCIs, opening a new application field for BCIs aimed at increasing the quality of life of motor-disabled people.

6.5 Limitations of the study

Despite the utility of our proposals, the present Doctoral Thesis has some limitations that should be discussed as well. The asynchronous management using thresholding has demonstrated its feasibility with target users. However, the technique has a strong dependence on the ERP classifier, which forces the supervisor to record additional non-control signals whenever the classifier is updated. That wrapper dependence was overcome by the MSE-based proposal but, unfortunately, it was not tested with motor-disabled patients. Moreover, the validation was made using an offline LOO procedure, using more training trials in each iteration than those that are generally used in practice. Therefore, the performance of the proposed MSE-based asynchronous method is expected to decay slightly in a real environment. Further endeavors should be focused on testing the method in an online application.

Even though meta-heuristics have demonstrated their utility in finding optimal channel sets for each user, they were only tested with HS subjects. Furthermore, since the purpose of the study was focused on the channel selection procedure, only basic feature extraction and classification algorithms were used. Whether these results could be improved or not using more sophisticated algorithms is still unknown. The proposed algorithms also entail high computational cost due to their wrapper nature, restricting their application after the calibration session. Moreover, hyperparameters were fixed according to the recommendations of the literature, but not optimized for each subject. It should be also noted that the 'III BCI Competition' database had more training trials than those that are usually recorded in practice.

Regarding the assistive applications, their asynchronous managements were composed by thresholding approaches. Owing to the results, MSE-based asynchrony exhibited a higher offline performance than the thresholding approach and thus, it is suggested that results could have been improved by using it. In both applications, patients requested a higher speed. Whether the classifier's performance and the speed could have been improved by using more robust classifiers or more training trials in the calibration session is an open question. Moreover, both applications lacked of despondency bypassing, causing a mistake to propagate further in some occasions. Regarding the social networking application, it should be highlighted that signal processing stage required a laptop to be executed, which favored the reliability but impaired the portability. In turn, it was tested with a heterogenous motor-disabled population, where a population with a certain disease

would have allowed to characterize its performance within that kind of patients. Likewise, in both applications the patients not only presented motor-disabilites, but also cognitive damage, which hindered the control of the BCIs.

The last limitation concerns the number of subjects that composed each database. Although this limitation is commonly present in BCI studies due to the difficulty to recruit motor-disabled patients, the sample size should be increased to favor the generalization of the results.

Chapter 7

Conclusions

During the previous chapters of this Doctoral Thesis it has been laid bare the common thread of the included studies: signal processing (i.e., asynchrony, channel selection) and assistive applications (i.e., web browser, social networking app) of P300-based BCIs toward a real use of these devices by motor-disabled people. These studies are intended to be a starting point to take the leap from laboratories and provide an actual use of BCIs in real environments. Among the signal processing approaches, it has been proposed two asynchronous methods to provide a comprehensive control of the BCIs, and a multi-objective meta-heuristic to select reduced but efficient customized channel sets. In fact, the asynchronous thresholding method was implemented in the assistive applications: the web browser and the mobile social networking app, whose feasibility has been assessed by target users.

In this chapter, the original contributions of this compendium of publications are highlighted in section 7.1. The next section 7.2 indicates the joint conclusions of these studies. Lastly, future endeavors related to this research are enumerated in section 7.3.

7.1 Contributions

The main contributions provided by this compendium of publications are listed next:

1) Development of a thresholding asynchronous method based on SW-LDA scores to monitor users' attention while using P300-based BCIs. Evaluation

of its feasibility with target users through its integration into two assistive applications (Martínez-Cagigal et al., 2017, 2019a).

2) Characterization of control and non-control state EEG signals during RCP stimulation through entropy-based measures. The analysis in terms of regularity and complexity led to the proposal of a MSE-based asynchronous method to avoid unintended selection of commands when using P300-based BCIs (Martínez-Cagigal et al., 2019b). To the best of our knowledge, this was the first time that asynchronous states were characterized through entropy measures and used to develop an asynchronous monitoring independent from the ERP classifier.

3) Comparison of single and multi-objective meta-heuristics to select appropriate sets of channels for each user. Although some existing algorithms were previously applied to this problem, to the best of our knowledge, this was the first time that their performances were compared, as well as evaluated with different paradigms apart from RCP (Martínez-Cagigal et al., 2020).

4) Proposal of DFGA, a novel multi-objective algorithm especially designed for the P300-based BCI channel selection problem. The algorithm demonstrated to reach similar or even higher performances than other approaches (MOPSO, NSGA-II, SPEA2), exhibiting a faster convergence (Martínez-Cagigal et al., 2020).

5) Design, development and evaluation of an asynchronous BCI web browser to assist motor-disabled people. The application was successfully tested with a population of 16 MS and 5 HS users, demonstrating its viability with target users (Martínez-Cagigal et al., 2017).

6) Design, development and evaluation of an asynchronous mobile BCI to control social networks (Twitter and Telegram) for those whose command of hands and fingers is limited. The system was assessed by a population of 18 motor-disabled and 10 HS, demonstrating its feasibility to bridge the accessibility gap in current smartphones (Martínez-Cagigal et al., 2019a).

7.2 Main conclusions

The analysis of the results (chapter 5) and its discussion (chapter 6) of the studies as a whole allows to draw the main conclusions of this Doctoral Thesis, which are enumerated below:

1) A practical P300-based BCI should implement an asynchronous management in order to avoid unintended command selections. This stage is essential to provide a comprehensive control of the system and avoid the dependence of supervisors, supporting the personal autonomy of the target users.

2) The integration of an asynchronous detection stage improves significantly the performance of the users in assistive BCI applications. The amount of selection mistakes is also reduced drastically.

3) EEG signals of users when attending to visual *oddball* stimuli are significantly more complex and irregular that those obtained when ignoring them. These differences allow to monitor their attention using entropy-based metrics.

4) Optimal channel sets depends strongly on the individual, showing a high inter-subject variability. Therefore, an optimization for each user is beneficial for the overall performance of the BCI, constituting a recommendable procedure if enough time is available after the calibration.

5) Discrete multi-objective meta-heuristics are suitable to find optimal sets in function of the number of electrodes, significantly outperforming the general rule-of-thumb of using eight channels for P300-based BCIs. A balanced combination of deterministic and stochastic techniques (e.g., DFGA) is also beneficial for their convergence.

6) Performances of motor-disabled users are significantly lower than those obtained by control subjects. Hence, assistive BCIs must be tested with target users in order to assure their feasibility in a real setup.

7) P300-based social networking through smartphones and web browsing have proven to be useful assistive applications for the motor-disabled. Their integration in the everyday life of dependent people is viable.

8) Opinions of the motor-disabled population reflect a general satisfaction with the assistive applications. Patients imagine themselves using P300-based BCIs in their daily life in the near future.

7.3 Future research lines

Several questions derived from this investigation may be the object of further endeavors to complement our findings, or even to take care of other topics that do

not belong to the scope of this Doctoral Thesis. In this regard, the most interesting future research lines are discussed next.

Concerning the MSE-based asynchronous approach, the combination of the proposed method with other complementary filter algorithms (e.g., spectral analysis) may improve its performance and constitutes an interesting future line of investigation. Moreover, the fact that the MSE-based algorithm does not depend on the ERP classifier leads us to encourage its assessment with a motor-disabled population to favor the generalization of the results. Studying whether the proposed feature is useful in characterizing other related aspects such as *BCI illiteracy* or inter-session variability would be interesting as well. Regarding the channel selection algorithms, the proposed meta-heuristics were evaluated using a simple processing pipeline. Future work could be focused on (i) improving the performance by complementing them with regularization techniques; or (ii) reducing their computational cost by integrating embedded strategies into the ERP classifiers. Another interesting future research could be the application of these meta-heuristics to feature selection problems.

With respect of the applications, the asynchronous management of both (web browser and social networking) employed the thresholding approach. Improvements could be made by integrating the MSE-based proposal, which is independent of the ERP classifier and avoids the recording of non-control data whenever is updated. One natural way to continue our investigation would be the assessment of these proposals with a larger motor-disabled database. Moreover, it would be interesting to focus on disabilities that do not entail cognitive problems, such as ALS or stroke; who constitute the fittest target users for BCIs. We also believe the general performance of the assistive applications may be improved by using more complex classification algorithms that could follow the inter-subject and inter-session variability of the ERPs; e.g., deep learning techniques. Finally, a long-term longitudinal validation of these applications could give insight into how socio-technological interactions and disease progression would affect the efficacy of the BCI system.

Appendix A

Papers included in this Doctoral Thesis

Asynchronous Control of P300-Based Brain–Computer Interfaces Using Sample Entropy

Víctor Martínez-Cagigal, Eduardo Santamaría-Vázquez, Roberto Hornero

Biomedical Engineering Group, E.T.S.I. Telecomunicación, University of Valladolid, Paseo de Belén 15, 47011, Valladolid, Spain

Entropy
Volume 25, No. 3, February 2019, Pages 230

Abstract

Brain–computer interfaces (BCI) have traditionally worked using synchronous paradigms. In recent years, much effort has been put into reaching asynchronous management, providing users with the ability to decide when a command should be selected. However, to the best of our knowledge, entropy metrics have not yet been explored. The present study has a twofold purpose: (i) to characterize both control and non-control states by examining the regularity of electroencephalography (EEG) signals; and (ii) to assess the efficacy of a scaled version of the sample entropy algorithm to provide asynchronous control for BCI systems. Ten healthy subjects participated in the study, who were asked to spell words through a visual oddball-based paradigm, attending (i.e., control) and ignoring (i.e., non-control) the stimuli. An optimization stage was performed for determining a common combination of hyperparameters for all subjects. Afterwards, these values were used to discern between both states using a linear classifier. Results show that control signals are more complex and irregular than non-control ones, reaching an average accuracy of 94.40% in classification. In conclusion, the present study demonstrates that the proposed framework is useful in monitoring the attention of a user, and granting the asynchrony of the BCI system.

Keywords: Sample entropy, multiscale entropy, brain–computer interfaces, asynchrony, event-related potentials, P300-evoked potentials, oddball paradigm

1. Introduction

Brain–computer interfaces (BCI) are able to detect users' intentions from brain signals and convert them into artificial commands that control an external device. BCI applications are intended to replace, restore, enhance, supplement, or improve the natural central-nervous-system activity of the user (Wolpaw and Wolpaw, 2012). Such purposes make BCI systems especially suited for improving the quality of life of motor-disabled people, reducing their dependence, and favoring their social and labor integration. These disabilities may be caused by traumas, neurodegenerative diseases, muscle disorders, or any illness that impairs the neural pathways that control muscles or the muscles themselves (Wolpaw et al., 2002). Although there are several ways to monitor the brain activity of a user, electroencephalography (EEG) is generally used due to its noninvasiveness, portability, and low cost. Therefore, electric brain activity is recorded by placing a set of electrodes on the user's scalp (Wolpaw et al., 2002).

Since a user's intentions are not directly reflected in the raw EEG signal, BCI systems rely on the processing of measurable changes related to cognitive tasks, known as control signals (Nicolas-Alonso and Gomez-Gil, 2012). Event-related potentials, such as P300 responses, are commonly used to assure the robustness of the system regardless of disability. P300-evoked potentials are the brain's natural responses to infrequent and significant stimuli, elicited approximately 300 ms after their onset (Wolpaw et al., 2002; Nicolas-Alonso and Gomez-Gil, 2012). Owing to their exogenous nature, previous training is not necessary, which makes a P300-based BCI suitable for any person who presents a certain degree of gaze control. In this sense, the row-col paradigm (RCP), a particularization of the oddball visual paradigm, is the most common setup to aid users in spelling words or commands (Farwell and Donchin, 1988). In this paradigm, a matrix containing alphanumeric characters or commands is displayed. Users just need to focus their attention on the desired command while the matrix's rows and columns randomly flash. Whenever the target's row or column is intensified, P300 potential is generated. Hence, the desired command can be determined by identifying when these potentials have been elicited (Farwell and Donchin, 1988; Wolpaw et al., 2002).

Email addresses: victor.martinez@gib.tel.uva.es (Víctor Martínez-Cagigal), eduardo.santamaria@gib.tel.uva.es (Eduardo Santamaría-Vázquez), robhor@tel.uva.es (Roberto Hornero)

DOI: https://doi.org/10.3390/e21030230

The RCP is a synchronous process. Due to continuous stimulation, the system makes a selection even if the user does not pay attention to the visual stimuli (Pinegger et al., 2015). In a real application, it is desirable that users voluntarily decide when they want to select a command and when they do not. For instance, if the purpose of the BCI system is to provide disabled users with an assistive tool to surf the Internet, the application should be able to detect if the user wants to select a navigation command or, by contrast, to calmly read a webpage or watch a video (Martínez-Cagigal et al., 2017). A conventional synchronous BCI could not monitor users' attention; thus, it continues selecting random commands while users ignore the visual stimulation. Therefore, the default synchronous mode of the RCP severely restricts the applicability of a BCI system in a real environment, requiring an external supervisor or the inclusion of a read-mode command that pauses the RCP for a fixed time. In order to overcome this limitation, the system should be able to discern between the control state (i.e., when users pay attention to the stimuli) and the non-control state (i.e., idle state, when users ignore the stimuli). In other words, the RCP-based system must become an asynchronous application. In recent years, several efforts have been made to achieve real asynchronous control (Pfurtscheller, 2010). Most related P300-based BCI studies rely on a threshold derived from classifiers' scores, which are expected to be higher in the control state than in the non-control state. These scores were obtained from support vector machines (SVM) (Zhang et al., 2008; Li et al., 2013; He et al., 2017) or linear discriminant analysis (LDA) (Aloise et al., 2011; Aref and Huggins, 2012; Breitwieser et al., 2016; Alcaide-Aguirre et al., 2017; Martínez-Cagigal et al., 2017; Tang et al., 2018; Martínez-Cagigal et al., 2019) classifiers using downsampled raw signals from the stimuli onset as features (Zhang et al., 2008; Panicker et al., 2010; Nicolas-Alonso and Gomez-Gil, 2012; Li et al., 2013; Pinegger et al., 2015; Martínez-Cagigal et al., 2017; Yu et al., 2017b; Martínez-Cagigal et al., 2018; Aydin et al., 2018; Martínez-Cagigal et al., 2019). Aydin et al. (2018) also used classifier labels instead of scores to design different criteria to identify the idle state. Other studies proposed spectral features to detect both states, such as relative powers (Panicker et al., 2010; Li et al., 2013; Ma and Qiu, 2018) or sums of spectral components (Pinegger et al., 2015). Among these complementary metrics, recent studies have proposed modified BCI frameworks. Panicker et al. (2010); Li et al. (2013) proposed novel asynchronous paradigms that involve steady-state visually evoked potentials (SSVEP) and P300 responses at the same time, using SSVEP to identify the idle state, and P300 responses to determine the desired commands in real time. Breitwieser et al. (2016) provided asynchronous control in a tactile-based BCI system to detect both steady-state somatosensory evoked potentials (SSSEP) and transient event-related potentials (tERP). Lastly, Yu et al. (2017b,a) presented a hybrid system that manages asynchronous control using motor imagery, while an RCP matrix controls the command selection.

Despite the recent interest in providing asynchronous control in RCP-based BCI systems, to the best of our knowledge, entropy metrics have not yet been explored. In this context, we hypothesize that different entropy metrics could provide insight into the dynamics of attended and nonattended EEG signals, providing complementary information to discern between both states. Particularly, multiscale entropy (MSE) based on sample entropy (SampEn) has demonstrated to be effective in estimating the complexity and regularity of physiological time series (Richman and Moorman, 2000; Costa et al., 2002, 2005; Humeau-Heurtier, 2015). Thus, differences between the regularity of control and non-control EEG signals could be expected to be found. Therefore, the present study has a twofold purpose: (i) to characterize control and non-control states by examining the regularity of EEG signals; and (ii) to assess the efficacy of a scaled version of SampEn to provide asynchronous control in P300-based BCI systems.

2. Materials and Methods

EEG signals show high intersubject variability and, thus, BCI systems must be optimized for each subject (Farwell and Donchin, 1988; Wolpaw et al., 2002; Nicolas-Alonso and Gomez-Gil, 2012; Martínez-Cagigal et al., 2017, 2018). The amplitude and latency of P300 responses have been demonstrated to vary depending on individual differences, such as age or personality, pharmacological aspects, or even clinical disorders (Picton, 1992). Therefore, channel and feature selection methods, as well as classifiers, are always optimized in the first session of each user. According to this rationale, classifiers of the present study are separately trained and tested, returning a final accuracy for each subject.

The methodological structure of the study is depicted in the flowchart of Figure 1. Once the dataset was registered and preprocessed, it was randomly divided into optimization (30%) and validation (70%) datasets. The optimization set was used to characterize the asynchronous states and find an optimal combination of the required hyperparameters that could work with all subjects. These global values were thereafter used to test the validation set for each user and assess the ability of the framework to discriminate between control and non-control states. Training and testing were employed under a leave-one-out (LOO) procedure, intended to provide a final accuracy for each user.

2.1. Dataset and Experimental Protocol

Ten control subjects (mean age 25.7 ± 3.09 years; 6 males, 4 females) were included in this study. All of them gave their informed written consent to participate. Subjects were asked to perform spelling tasks using a 6×6

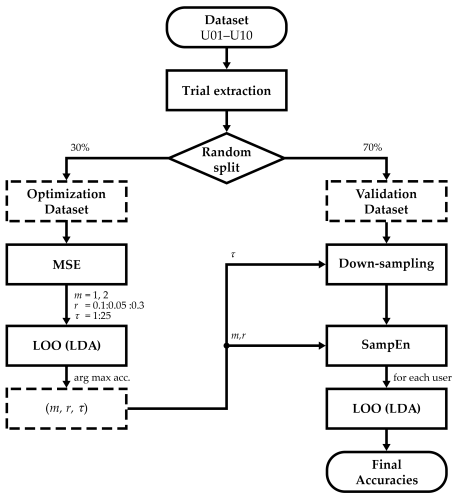


Figure 1: Methodological flowchart of the study. Once trials were extracted, the dataset was divided into optimization and validation sets. The former was intended to optimize a global combination of hyperparameters m, r, and τ; in the latter, these values were applied to compute the final accuracy of each user.

RCP matrix in two different sessions, shown in the Figure 2a. In the RCP paradigm, the matrix's rows and columns randomly flash (Farwell and Donchin, 1988). Users, who were asked to stare at the desired command, elicited P300 responses when the row and the column that contained that command were illuminated. Therefore, the desired command could be determined by identifying these responses (Farwell and Donchin, 1988; Wolpaw et al., 2002). In order to favor their concentration, users were also asked to count how many times the desired command flashed. For each user, a total of 120 characters were spelled. Half of them were recorded following the aforementioned protocol, intended to get the signal in the control state. For the other half, users were asked to read a text while ignoring the flashings. Hence, these characters were intended to record the non-control state. Note that a character comprised 15 sequences (i.e., repetitions) of flashings, where a sequence comprises all flashes that are required to highlight each row and column of the matrix. Each flashing lasted 75 ms, followed by an interstimuli interval of 100 ms. EEG signals were recorded using a g.USBamp amplifier (g.Tec, Austria) with a sampling rate of 256 Hz. In all, 16 active electrodes were placed on Fz, F3, F4, Cz, C3, C4, CPz, Pz, P3, P4, POz, PO3, PO4, PO7, PO8, and Oz, using Fpz as a ground and the earlobe as a reference according to the International 10–20 System distribution (Jasper, 1958). Since P300 responses are thought to be more prominent over the visual cortex and related with cognitive processing, electrodes were mainly placed on the occipital and parietal lobes (Wolpaw et al., 2002; Nicolas-Alonso and Gomez-Gil, 2012).

As a preprocessing stage, a band-pass filter in the range of 0.1–30 Hz and a common average reference (CAR) spatial filter were applied to the raw signals (Wolpaw et al., 2002; Martínez-Cagigal et al., 2017, 2018). Afterward, trials were extracted from the EEG signals for each channel following the procedure that is depicted in Figure 2b. As can be seen, each trial integrates the signal from the first sample to the last onset that belongs to the maximum considered sequence. For instance, the i-th trial comprises the raw signal of all electrodes since the very first recording sample of the character until the end of the i-th sequence. Then, the dataset was randomly split up into optimization (30%) and validation (70%) sets. These ratios were maintained for each user, resulting in a total of 36 characters for the optimization set and 84 characters for the validation set per user. It is noteworthy that both sets were also balanced, including the same number of control and non-control characters of each user.

2.2. Optimization Stage

The optimization stage was intended to find a global combination of hyperparameters that favor the discrimination between control and non-control states for all users. To this end, features were first extracted by means of MSE, and then classified with an LDA following a LOO procedure. As indicated in Figure 1, the combination of parameters was finally selected under a criterion of maximum performance.

MSE is a well-known nonlinear method that estimates the complexity of a signal according to entropy changes along multiple time scales (Costa et al., 2002). The algorithm sequentially computes the entropy of a coarse-grained version of the original signal, providing information about its dynamical structure (Costa et al., 2002, 2005). If MSE is applied on two different time series, and one of them provides higher entropy values for most scales, it is considered to be more complex (Costa et al., 2002, 2005). Typically, the τ-th scaled coarse-grained signal is obtained by averaging the samples of the time series inside consecutive but nonoverlapped segments of length N/τ, where N denotes the length of the signal (Costa et al., 2002). However, it was shown that this procedure may cause aliasing and, thus, spurious components in the low-frequency range (Valencia et al., 2009; Humeau-Heurtier, 2015). In order to overcome this limitation, we decimated the original signal by a factor of τ. That is, high frequencies were reduced with a low-pass least-squares linear-phase FIR filter, followed by a downsampling procedure that only kept every τ-th sample (Valencia et al., 2009; Humeau-Heurtier, 2015). Therefore, the MSE algorithm computes the entropy of each signal as a function of τ

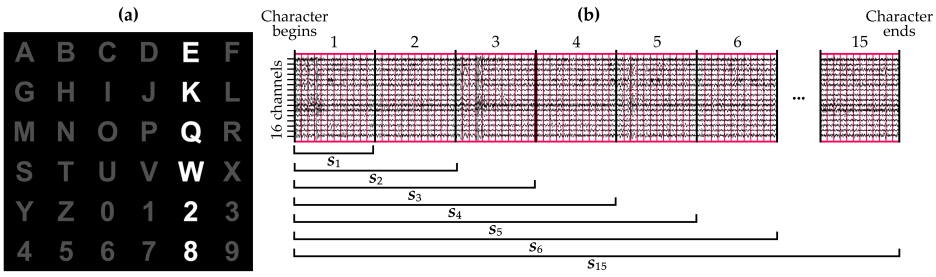


Figure 2: (**a**) Row-col paradigm matrix employed in this study. Currently, the fifth column is being flashed; (**b**) Trial extraction procedure of a single character in function of the number of sequences. Considering the i-th sequence, trial $\boldsymbol{s}_i$ is composed of the signal from the first sample to the last onset of the i-th sequence. Therefore, a total of 15 trials were extracted for each character.

from the original time series (i.e., $\tau = 1$), to the highest considered scale (i.e., $\tau = 25$) (Humeau-Heurtier, 2015).

SampEn is a single-scale entropy measure that estimates the irregularity of one-dimensional temporal signals, assigning higher values to series that show larger degrees of disorder (Richman and Moorman, 2000). Compared to the approximate entropy algorithm, SampEn eliminates the inherent bias caused by self-matching and provides a result less dependent on signal length (Richman and Moorman, 2000). For this reason, SampEn has been widely used to compute the MSE and its variants (Humeau-Heurtier, 2015). Briefly, the algorithm provides a conditional probability measure that quantifies the likelihood that a template of m consecutive samples, which already matches another sequence, still matches it if their lengths are increased in one sample (Humeau-Heurtier, 2015). Therefore, SampEn is defined as:

$$SampEn(m, r, N) = \lim_{N \to \infty} -ln \frac{A^m(r, N)}{B^m(r, N)}, \qquad (1)$$

where m is the embedding dimension, r is the tolerance factor, N is the length of the signal, and $A^m(r, N)$ and $B^m(r, N)$ are the probabilities of template matching for $m + 1$ and m points, respectively. Considering a time series $\boldsymbol{x} = [x_1, x_2, \ldots, x_N]$, where template vectors of length m are defined as $\boldsymbol{x}_m(i) = [x_i, x_{i+1}, \ldots x_{i+m-1}]$, a match between two templates $\boldsymbol{x}_m(i)$ and $\boldsymbol{x}_m(j)$ occur if the distance between them is less than a certain tolerance value: $d[\boldsymbol{x}_m(i), \boldsymbol{x}_m(j)] < R$. Although there are a variety of distance measures, Chebyshev distance is commonly used (Richman and Moorman, 2000). Moreover, tolerance is used to be dependent of the standard deviation of the signal (i.e., $R = r \cdot \sigma_{\boldsymbol{x}}$) (Richman and Moorman, 2000; Humeau-Heurtier, 2015). In practice, SampEn is estimated as follows:

$$SampEn(m, r, N) = -ln\Big(\frac{N - m + 1}{N - m - 1} \cdot \frac{A}{B}\Big), \qquad (2)$$

where A and B are the total number of templates of lengths $m + 1$ and m that meet the distance criterion for each different combination of i and j, given $i \neq j$, respectively. Since the total number of possible templates of lengths $m + 1$ and m along the signal are $N - m + 1$ and $N - m - 1$, respectively; normalization is also applied to correct the estimation. As a result of the approximation of Equation (1), the variance of the entropy estimator grows as the length of the signal decreases (Humeau-Heurtier, 2015). Therefore, the longer the signal length, the more reliable the outcome is. As a general rule of thumb, the estimation of SampEn is considered accurate if $N \geq 10^m$ (Richman and Moorman, 2000; Humeau-Heurtier, 2015).

MSE using a SampEn estimator was then applied to the optimization dataset. Hyperparameters were varied according to common ranges widely used in physiological signals: embedding dimensions $m = 1, 2$; tolerances r from 0.1 to 0.3 in steps of 0.05; and scales τ from 1 to 25 (Richman and Moorman, 2000). Scales that did not meet the Richman & Moorman criterion (i.e., $N \geq 10^m$) were not computed (Richman and Moorman, 2000). Since entropies should be estimated in one-dimensional signals, MSE was calculated for each channel, returning a final value per channel and trial. Note that trials were extracted following the procedure described in Section 2.1, computing the MSE using different number of RCP sequences, from 1 to 15. Figure 3 depicts the MSE results of the user U05 for illustrative purposes.

In order to determine a common optimal combination of τ, m, and r for all users, an LOO procedure was performed. LOO cross-validation is a deterministic technique that estimates how the results of a statistical model generalize to an independent dataset (Witten and Frank, 2011). The algorithm sequentially classifies an observation with a model trained with the remaining ones. This process is repeated until all observations have been tested, returning the average of the prediction outcomes as an estimation of the accuracy (Witten and Frank, 2011). In this case, the LOO procedure integrated an LDA that

classified control versus non-control observations, where MSE results of each channel were included as features. The accuracies for all trials, $s_1, s_2, \ldots, s_{15}$, were averaged in order to get a single accuracy value for each combination of τ, m, and r. Lastly, the combination of hyperparameters that reached maximum accuracy was thereafter considered optimal. Owing to the mix of users that composes the optimization dataset, the optimal m, r, and τ are expected to work properly regardless of subject.

2.3. Validation Stage

The validation stage was intended to assess the performance of the proposed framework to determine the state of the user and achieve asynchronous control of the system. As can be noticed, since MSE was not computed to consider any geometric feature of the curve, but to determine an optimal scale τ, there is no point in calculating the MSE in the validation dataset. Instead, validation signals for each user are first downsampled to optimal scale τ. Afterward, features are the SampEn outcomes of each channel using optimal m and r hyperparameters. An LDA-based LOO procedure is finally used to estimate the accuracy of the classification per user and sequence.

3. Results

Optimization results are depicted in Figure 4. As can be seen, the estimated accuracies show a decreasing tendency as the scale increases regardless of embedding dimension. According to the maximum-accuracy criterion, the optimal combination of hyperparameters was found to be $m = 1$, $r = 0.3$, and $\tau = 2$. Figure 5 depicts the spatial distribution of the significant differences that were found between control and non-control SampEn features in the optimization dataset (Wilcoxon signed rank test), using the aforementioned optimal parameters. It is noteworthy that the Benjamini–Hochberg False Discovery Rate (FDR) correction was applied to counteract the problem of multiple comparisons (i.e., 16 channels) (Benjamini and Hochberg, 1995). As shown, significant differences were mainly found in prefrontal and occipital electrodes.

The results of the validation stage are displayed in Table 1 and Figure 6. The proposed framework reached a mean accuracy of $94.40\% \pm 2.81\%$ across subjects for 15 sequences. Figure 6 depicts the cumulative testing accuracies (control vs. non-control) as the number of sequence increases for each subject. As can be seen, users generally showed an improvement in performance as more sequences are considered, reaching more than 90% of accuracy in every case. In order to guarantee the application of the proposed framework in real time, computational cost analysis is shown in the Table 2, which details the required time to compute the SampEn algorithm using different numbers of sequences. Analysis was made using an Intel Core i7-7700 CPU @ 3.60GHz (32 GB RAM, Windows 10, MATLAB®2018a), performing an average of 1000 iterations of the algorithm.

4. Discussion

Significant differences were found between control and non-control states using features derived from MSE and SampEn. Since the depicted behavior of Figure 3 is representative of all subjects, SampEn values of control states were slightly higher than those obtained in non-control states. Moreover, this behavior is almost constant as scales increase (i.e., amount of decimation). The MSE values of both states show an increasing trend until $\tau = 4$, steadying themselves after that point. On the one hand, this tendency implies that attending to an RCP paradigm produces more irregular signals than ignoring the stimuli (Richman and Moorman, 2000). On the other hand, although both states show a similar response to dynamical changes in different scales, control signals present a steeper slope. Therefore, control-state signals can be considered more complex than non-control ones because they are more irregular in most scales (Costa et al., 2002, 2005). It is also noteworthy that SampEn values of nonattending signals become more unstable as the scale increases, raising the standard deviation. By contrast, attending signals are generally more defined, showing smaller values of standard deviation.

Regarding the optimization stage, it is noteworthy that the performance of the method depends on the hyperparameters. Although MSE values do not seem to be affected by tolerance, the embedding dimension and the scale play an important role in the proposed framework. As can be seen in Figure 4, performance showed a decreasing tendency as τ increased, regardless of the value of r. As aforementioned, the standard deviation of non-control MSE values increases with τ, while control MSE values remain almost constant. Hence, the decrease in performance for high scales is expected. Although accuracy values when $m = 1$ are not appreciably affected by r, performance decays as r decreases when $m = 2$. This behavior is also expected according to the SampEn algorithm, since higher tolerance values increment the probability of finding template matchings and, thus, increasing variability between different runs of the LOO procedure. In summary, the optimal embedding dimension and tolerance parameters were found to be $m = 1$ and $r = 0.3$, respectively, in accordance with previous studies that used physiological signals (Richman and Moorman, 2000). Concerning the optimal $\tau = 2$ scale, it is equivalent to reducing the sampling rate of the EEG signal by half before applying the SampEn algorithm. This procedure can be addressed as a feature-extraction stage that is common in P300-based BCI studies (Zhang et al., 2008; Panicker et al., 2010; Nicolas-Alonso and Gomez-Gil, 2012; Li et al., 2013; Pinegger et al., 2015; Martínez-Cagigal et al.,

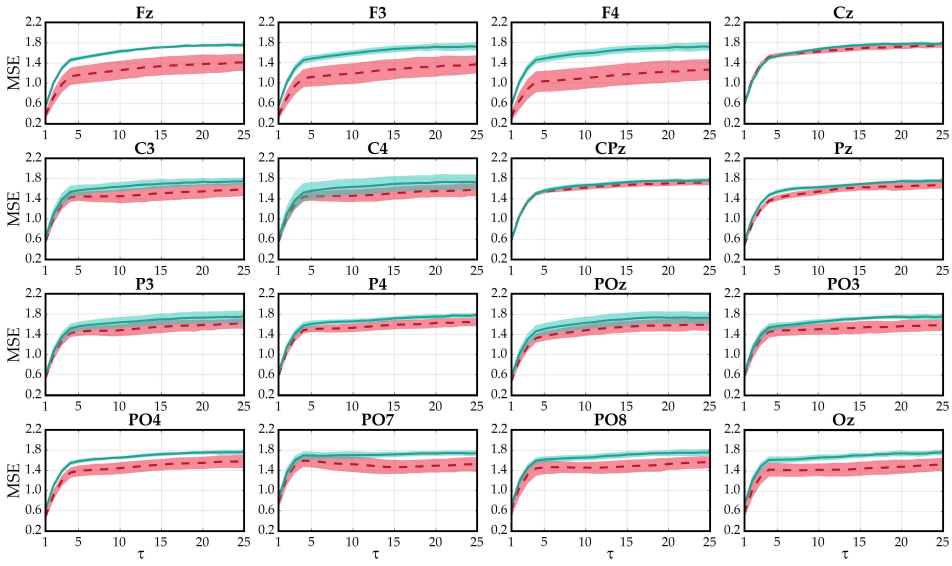


Figure 3: Multiscale sample entropy values from the optimization dataset corresponding to U05 across channels. Solid lines indicate the average values for control (blue) and non-control (red) trials, whereas shaded areas indicate standard deviation. Embedding dimension and tolerance parameters were fixed to $m = 1$ and $r = 0.3$, respectively.

2017; Yu et al., 2017b; Martínez-Cagigal et al., 2018; Aydin et al., 2018; Martínez-Cagigal et al., 2019).

In this context, the estimation of SampEn could be considered accurate when signal length is greater than ten to the power of the embedding dimension (i.e., $N \geq 10^m$) (Richman and Moorman, 2000; Humeau-Heurtier, 2015). According to Figure 2, signal length depends on the number of sequences that are considered, as well as on the amount of decimation. Since this limitation takes into account the amount of raw samples, the maximum number of scales that can be computed in a reliable way are thus limited by the number of sequences, the sampling rate, the stimuli duration, the number of commands and, in general, by any parameter that affects the duration of a character trial. In a P300-based BCI common setup, this constraint is not usually present for a high number of sequences (i.e., N_s), but it is recommended to compute the maximal scale in each situation. In our study, the entire number of 25 scales could be computed if $N_s > 4$, reaching a maximum of four scales using only one sequence. Owing to fixing the optimal scale to $\tau = 2$, the constraint did not even limit the number of sequences in our case.

Topographic results show significant differences for almost all users between the entropy values of control and non-control states, mainly in the prefrontal lobe. The prefrontal cortex is commonly associated with planning complex cognitive behavior, personality expression, decision making, and selective attention (Lebedev et al., 2004). The latter is consistent with the oddball task, which implies a constant attention of the user to identify the target stimuli among other background stimuli (Nicolas-Alonso and Gomez-Gil, 2012). In fact, it was demonstrated that visual oddball tasks produce hemodynamic changes in the dorsolateral prefrontal cortex, associated with the mapping of stimuli to responses (e.g., response strategies) (Huettel and McCarthy, 2004). Moreover, a recent study suggested that complex processes such as memory, attention, or decision making are linked to the elicitation of the P300 component, which could be modulated by frequency dynamics (Bojorges-Valdez and Yanez-Suarez, 2018). There are also slight differences in the occipital lobe, which comprises most of the anatomical region of the visual cortex. Neurons of the primary visual cortex fire action potentials when visual stimuli appear in the receptive field (Goodale et al., 1982). It is therefore expected that a higher number of neurons are activated in the control state, when a user not only perceives the target stimuli, but also repetitive flickering stimuli. The task elicits P300-evoked potentials in the parietal cortex when target stimuli are processed (Wolpaw et al., 2002; Nicolas-Alonso and Gomez-Gil, 2012). How-

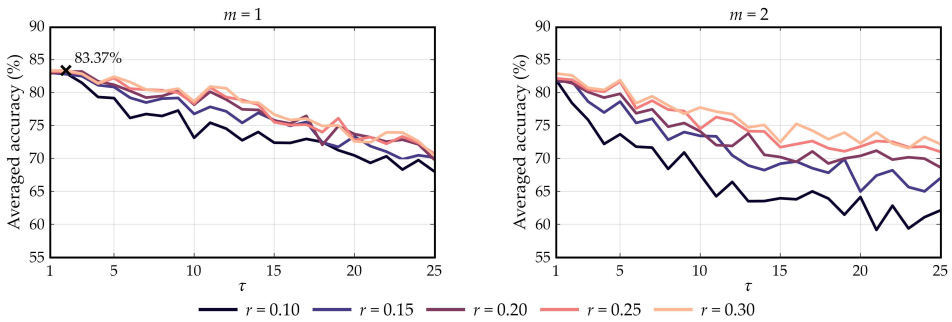


Figure 4: Accuracy results of the optimization stage in function of different values of embedding dimension m, tolerance r, and scale τ. Optimal combination of hyperparameters is marked with a cross, which corresponds to $m = 1$, $r = 0.3$, and $\tau = 2$.

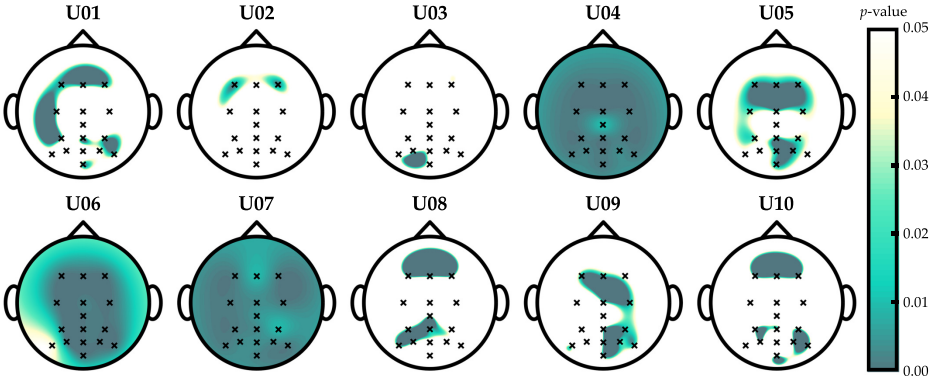


Figure 5: Wilcoxon signed-rank test p-values that show significant differences (i.e., from 0 to 0.05) between control and non-control SampEn features in the optimization dataset. Hyperparameters were fixed to their optimal values. Note that p-values were adjusted using the Benjamini–Hochberg False Discovery Rate (FDR) step-up procedure.

ever, since we extract features using the entire raw control EEG signal, P300 are surpassed by nontarget stimuli. Recent studies suggest that peripheral flickering stimuli in the RCP task produce SSVEP responses (Panicker et al., 2010; Pinegger et al., 2015; Ma and Qiu, 2018), which propagate from occipital to prefrontal electrodes (Srinivasan et al., 2006). Note that these topographic results measure significant differences between the irregularity of control- and non-control-state EEG signals. According to previous analysis, attending to a RCP task should activate a greater number of neurons than ignoring the stimuli, spreading electrical activity across the frequency spectrum. Therefore, entropy measures follow that spectral activation, increasing the irregularity of the control signals.

One of the most crucial obstacles of BCI systems is to find methods that can be applied in real time. In relation to this, we consider important to analyze the potential of the proposed framework to determine the asynchronous state upon which a character is selected. As indicated in Table 2, the maximal computational time of the SampEn algorithm is approximately 197 ms using 15 sequences. Since most P300-based BCI studies use pauses of at least two seconds after each character, the computational cost of the proposed framework is perfectly acceptable (Schalk et al., 2004; Martínez-Cagigal et al., 2017, 2018, 2019).

Concerning the validation stage, Figure 6 and Table 1 show an increasing tendency of the final accuracies for all subjects as the number of sequences increases. Therefore, it is clear that the proposed asynchrony approach is dependent on the length of the signals, reaching an average accuracy of 94.40% for all subjects using 15 sequences.

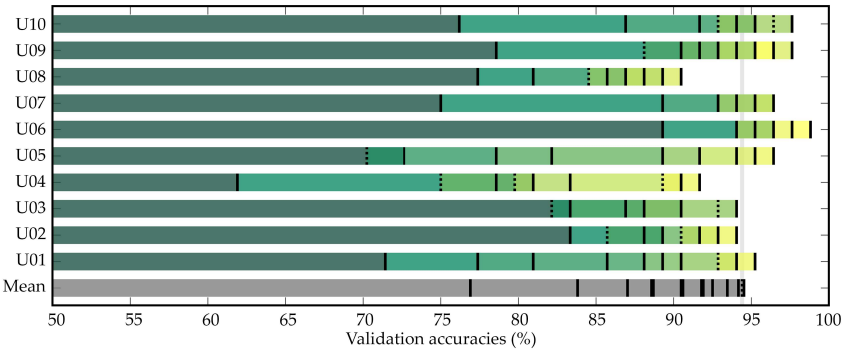


Figure 6: Cumulative testing accuracies (control vs. non-control) as sequences increase for each subject. Lines indicate the number of sequences, where a solid line implies an increase and a dashed line implies a decrease of accuracy.

Table 1: Testing accuracies of control vs. non-control states for each subject in function of the number of sequences.

N_s	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
U01	71.43%	77.38%	80.95%	85.71%	88.10%	88.10%	89.29%	90.48%	94.05%	94.05%	92.86%	92.86%	94.05%	95.24%	94.05%
U02	83.33%	88.10%	89.29%	85.71%	89.29%	89.29%	91.67%	91.67%	91.67%	90.48%	92.86%	91.67%	92.86%	94.05%	92.86%
U03	83.33%	82.14%	88.10%	83.33%	86.90%	90.48%	88.10%	90.48%	94.05%	92.86%	92.86%	92.86%	92.86%	92.86%	92.86%
U04	61.90%	78.57%	80.95%	75.00%	75.00%	75.00%	80.95%	80.95%	79.76%	80.95%	83.33%	90.48%	89.29%	91.67%	91.67%
U05	72.62%	70.24%	72.62%	78.57%	78.57%	82.14%	89.29%	89.29%	91.67%	91.67%	91.67%	94.05%	95.24%	96.43%	96.43%
U06	89.29%	94.05%	96.43%	96.43%	95.24%	94.05%	96.43%	95.24%	94.05%	95.24%	96.43%	96.43%	96.43%	97.62%	98.81%
U07	75.00%	89.29%	92.86%	95.24%	96.43%	96.43%	95.24%	95.24%	92.86%	94.05%	95.24%	95.24%	96.43%	96.43%	95.24%
U08	77.38%	80.95%	85.71%	86.90%	86.90%	88.10%	86.90%	84.52%	89.29%	89.29%	86.90%	89.29%	90.48%	89.29%	89.29%
U09	78.57%	90.48%	91.67%	88.10%	94.05%	90.48%	92.86%	95.24%	95.24%	92.86%	95.24%	95.24%	97.62%	95.24%	96.43%
U10	76.19%	86.90%	91.67%	95.24%	95.24%	92.86%	94.05%	92.86%	95.24%	97.62%	97.62%	96.43%	96.43%	96.43%	96.43%
Mean	**76.90%**	**83.81%**	**87.02%**	**87.02%**	**88.57%**	**88.69%**	**90.48%**	**90.60%**	**91.79%**	**91.90%**	**92.50%**	**93.45%**	**94.17%**	**94.52%**	**94.40%**
SD	7.58%	7.23%	7.11%	7.13%	7.23%	6.18%	4.59%	4.74%	4.61%	4.52%	4.38%	2.46%	2.77%	2.58%	2.81%

N_s indicates number of sequences.

In particular, all subjects except U04 and U08 reached more than 90% accuracy using nine sequences. Furthermore, U06, U07, and U09 were even able to reach it using only three sequences. Even though the increasing tendency is clear for all subjects, the slope appreciably varies among them. Some users present a sequential increase (e.g., U01, U05, U08, U09), while others reach a standstill (e.g., U03, U10). These results reinforce the fact that it is important to perform individual calibrations and separately optimize BCI applications to each subject (Wolpaw et al., 2002; Nicolas-Alonso and Gomez-Gil, 2012; Martínez-Cagigal et al., 2018).

Table 3 depicts a comparison between previous asynchronous P300-based state-of-the-art applications. As shown, most of them follow a thresholding approach to discern between control and non-control states (Zhang et al., 2008; Aloise et al., 2011; Li et al., 2013; Pinegger et al., 2015; Breitwieser et al., 2016; Martínez-Cagigal et al., 2017; Aref and Huggins, 2012; Alcaide-Aguirre et al., 2017; Ma and Qiu, 2018; Aydin et al., 2018; Tang et al., 2018; Martínez-Cagigal et al., 2019). These thresholds are usually derived from receiver operating characteristic (ROC) curves that were fed using output scores of SVM (Zhang et al., 2008; Li et al., 2013) or LDA (Aloise et al., 2011; Pinegger et al., 2015; Breitwieser et al., 2016; Martínez-Cagigal et al., 2017; Tang et al., 2018; Martínez-Cagigal et al., 2019) classifiers. Note that these classifiers use downsampled raw signals from the stimuli onset as input features (Zhang et al., 2008; Panicker et al., 2010; Nicolas-Alonso and Gomez-Gil, 2012; Li et al., 2013; Pinegger et al., 2015; Martínez-Cagigal et al., 2017; Yu et al., 2017b; Martínez-Cagigal et al., 2018; Aydin et al., 2018; Martínez-Cagigal et al., 2019). Since they were trained in a calibration session to detect P300 responses, these studies hypothesize that output scores of non-control characters are lower than those spelled in the control state. Therefore, the classifier that is intended to detect the P300 responses is also intended to discern between both asynchronous states. Notwithstanding their usefulness as computationally simple solutions, these approaches entail a clear drawback. Owing to the high intersession variability of the EEG signals, classifier weights should be updated from time to time to assure suitable performance (Wolpaw et al., 2002;

Table 2: Computational cost in milliseconds of the sample entropy algorithm in function of the number of sequences.

N_s	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Mean	0.82	3.46	8.24	14.64	22.57	32.51	43.66	54.92	69.70	86.33	104.87	125.24	146.41	170.58	196.78
SD	0.99	0.28	0.82	1.03	1.40	2.00	3.10	3.30	3.84	4.69	5.56	5.96	6.50	7.20	8.64

N_s indicates the number of sequences. These results are obtained after running the sample entropy algorithm 1000 times.

Nicolas-Alonso and Gomez-Gil, 2012; He et al., 2017; Martínez-Cagigal et al., 2017, 2019). Since threshold values depend on classifier scores, they are no longer useful if these weights are modified. Hence, additional control and non-control characters should be recorded in order to update the thresholds, which would entail a great amount of time. Other approaches add complementary spectral features (Pinegger et al., 2015; Ma and Qiu, 2018) or implement hybrid paradigms (Panicker et al., 2010; Li et al., 2013; Yu et al., 2017b) to develop filter methods that are independent of the P300 classifier. Some of the hybrid paradigms superimpose the RCP oddball technique, intended to generate P300 responses, with a flickering visual effect, intended to generate SSVEPs when users are paying attention to the visual stimuli (Panicker et al., 2010; Li et al., 2013). Therefore, asynchrony is handled by the detection of SSVEPs using relative powers: control state if SSVEPs are present, non-control state if SSVEPs are missing (Panicker et al., 2010; Li et al., 2013). Pinegger et al. (2015); Ma and Qiu (2018) also used SSVEP detection techniques to reach asynchronous control, but their approach is utterly different. By contrast, they hypothesized that inherent RCP flashings also generate residual SSVEP components when the stimuli are displayed using a constant rate. These components were identified in the frequency spectrum, providing supplementary features to the LDA scores (Pinegger et al., 2015). Finally, it is also worthy to mention the contribution of Yu et al. (2017b), who implemented a hybrid approach to reach a semiasynchronous BCI application. Users activated the RCP flashings by regulating their cortical activity through motor imagery. However, stopping RCP was handled by a "stop" command, which increases the required time to manage the asynchrony and makes the system more demanding. Since the vast majority of these previous studies were intended to provide an asynchronous assistive application, instead of just evaluating a novel method to reach asynchronous control, the provided accuracies reflect the final performance of the system. In other words, results depict the performance of the system to predict correct characters, while ignoring those than are considered non-control. Unfortunately, control versus non-control accuracies are not reported and, thus, quantitative and statistical comparisons cannot be performed with the present study. Despite this issue, it is noteworthy that, to the best of our knowledge, there are no studies that have previously investigated the ability of entropy-based features to discern between both asynchronous states. Moreover, since our approach is independent of classifier, weights updates do not affect asynchronous management, avoiding the need to record extra EEG signals (Zhang et al., 2008; Aloise et al., 2011; Li et al., 2013; Pinegger et al., 2015; Breitwieser et al., 2016; Martínez-Cagigal et al., 2017; Aref and Huggins, 2012; Alcaide-Aguirre et al., 2017; Aydin et al., 2018; Tang et al., 2018; Martínez-Cagigal et al., 2019). We also believe that further endeavors could be aimed at complementing our proposed entropy features with SSVEP-based ones, which could presumably improve the final performance of asynchronous P300-based BCI systems (Pinegger et al., 2015; Ma and Qiu, 2018).

Owing to these outcomes, several insightful implications can be derived. First, it was demonstrated that a scaled version of SampEn can follow the dynamic changes of control and non-control EEG signals, providing a useful tool to monitor the attention of the user. Furthermore, the proposed framework is not only able to work in real time for P300-based BCI systems, but also may be considered as a filter method. In other words, the metric is independent of the classifier that determines the selected command, in contrast with previous approaches (Zhang et al., 2008; Aloise et al., 2011; Li et al., 2013; Pinegger et al., 2015; Breitwieser et al., 2016; Martínez-Cagigal et al., 2017; Aref and Huggins, 2012; Alcaide-Aguirre et al., 2017; Ma and Qiu, 2018; Aydin et al., 2018; Tang et al., 2018; Martínez-Cagigal et al., 2019). Since our proposal does not rely on the classifier's scores, the command-oriented classifier can be updated without requiring a further training of the asynchrony method. Moreover, both states were also analyzed in this study, showing that control-state signals are more irregular and complex than non-control ones. Finally, a combination of user-independent hyperparameters were determined. To summarize, it was demonstrated that the proposed SampEn-based framework is suitable for providing asynchronous control in P300-based BCI systems.

In spite of these results, the present study has several limitations. The proposed framework only employed temporal features derived from the SampEn algorithm to classify between control and non-control states. The performance of this approach could be extended in the future by integrating complementary spectral features in order to improve its performance (Pinegger et al., 2015; Panicker et al., 2010; Li et al., 2013; Breitwieser et al., 2016). It is also noteworthy that the global combination of hyperparameters was defined using 10 control subjects who are not the target users of BCI systems. A future endeavor should be aimed at incrementing the database with both

Table 3: Comparison between previous asynchronous P300-based brain–computer interface (BCI) applications.

Study	Control Signal	Experimental Paradigm	Asynchrony Technique	No. Subjects
Zhang et al. (2008)	P300	Single cell	ROC thresholding using SVM scores	4 CS
Panicker et al. (2010)	P300 and SSVEP	Hybrid: RCP-based	Detection of SSVEPs using relative peak amplitude in PSD	10 CS
Aloise et al. (2011)	P300	RCP	ROC thresholding using LDA scores	11 CS
Li et al. (2013)	P300 & SSVEP	Hybrid: oddball & SSVEP	ROC thresholding using SVM scores (P300) and relative powers (SSVEP)	8 CS
Pinegger et al. (2015)	P300	RCP	Thresholding using LDA scores and sum of spectral components	10 CS
Breitwieser et al. (2016)	P300 and SSSEP	Hybrid: tactile & oddball	Thresholding using multi-class LDA	14 CS
Martínez-Cagigal et al. (2017)	P300	RCP	ROC thresholding using LDA scores	5 CS, 16 MS
He et al. (2017)	P300	RCP	Combination of two different SVM	8 CS
Yu et al. (2017b,a)	P300 and MI	MI monitoring & RCP	MI signal activates the RCP	11 CS, 8 CS
Aref and Huggins (2012); Alcaide-Aguirre et al. (2017)	P300	RCP	Certainty algorithm: t-test over LDA scores	11 CS, 19 CP
Ma and Qiu (2018)	P300	RCP	ROC thresholding using relative powers	4 CS
Aydin et al. (2018)	P300	Hex-o-Spell	ROC thresholding using classifier labels	10 CS
Tang et al. (2018)	P300	RCP	ROC thresholding using LDA scores	4 CS
Martínez-Cagigal et al. (2019)	P300	RCP	ROC thresholding using LDA scores	18 CS, 10 MD
Present study	**P300**	**RCP**	**LDA classification using SampEn features**	**10 CS**

SSVEP: steady-state visual evoked potentials, SSSEP: somatosensory evoked potentials, MI: motor imagery, RCP: row-col paradigm, ROC: receiver operating characteristic, SVM: support vector machines, PSD: power spectral density, LDA: linear discriminant analysis, SampEn: sample entropy, CS: control subjects, MS: multiple sclerosis, CP: cerebral palsy, MD: motor-disabled.

control and motor-disabled users in order to improve the generalization of these results. Furthermore, the variability of the optimal hyperparameters was not addressed in this study. Finally, it should be noted that the validation stage was applied under an LOO procedure. Although this method is excellent to estimate the performance of a statistical model, it requires more training trials in each iteration than those that are commonly used in practice. Moreover, owing to the limited number of subjects and characters in the database, optimization could not be performed using different users than in the validation procedure.

5. Conclusions

In this study, differences between control and non-control signal was analyzed using entropy metrics. Furthermore, a method to discern between both states and provide an asynchronous control of a P300-based BCI has been proposed. Dataset was composed of the EEG signals of ten healthy subjects who were asked to perform spelling tasks in a row-col paradigm, attending and ignoring the stimuli. Signals were then subdivided into optimization and validation sets. The former was used to determine a common optimal combination of hyperparameters by applying MSE features in a LOO procedure. These parameters were thereafter fixed at $m = 1$, $r = 0.3$, and $\tau = 2$ for all subjects. Then, the latter was used to test the ability of a scaled version of SampEn to characterize both states. Multiscale analysis results showed that control signals are more irregular and complex than non-control ones, regardless of scale. These features were also demonstrated to be suitable for classifying both states, reaching an average accuracy of 94.40%. From the experimental outcomes of this exploratory research, we conclude that: (i) MSE measures could follow the dynamic changes of control and non-control signals; (ii) the optimal combination of hyperparameters favors the discrimination between both states for all control subjects; (iii) the proposed framework has the potential to provide asynchronous control with high accuracies; and (iv) the computational cost of the method is negligible, reaching real-time processing.

Acknowledgments

This study was partially funded by projects DPI2017-84280-R of Ministerio de Ciencia, Innovación y Universidades and the European Regional Development Fund (FEDER), and the project "Análisis y correlación entre el genoma completo y la actividad cerebral para la ayuda en el diagnóstico de la enfermedad de Alzheimer" (Inter-regional co-operation program VA Spain–Portugal POCTEP 2014-2020) of the European Commission and FEDER. V.M.-C. was in receipt of a PIF-UVa grant of the University of Valladolid.

References

Alcaide-Aguirre, R. E., Warschausky, S. A., Brown, D., Aref, A., Huggins, J. E., 2017. Asynchronous brain–computer interface for cognitive assessment in people with cerebral palsy. Journal of Neural Engineering 14 (066001), 1–10.

Aloise, F., Schettini, F., Aricò, P., Leotta, F., Salinari, S., Mattia, D., Babiloni, F., Cincotti, F., 2011. P300-based brain–computer interface for environmental control: an asynchronous approach. Journal of Neural Engineering 8 (2), 025025.

Aref, A., Huggins, J., 2012. The P300-certainty algorithm: improving accuracy by withholding erroneous selections. In: EEG and Clinical Neuroscience Society Conference 2012.

Aydin, E. A., Bay, O. F., Guler, I., 2018. P300-Based Asynchronous Brain Computer Interface for Environmental Control System. IEEE Journal of Biomedical and Health Informatics 22 (3), 653–663.

Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate : A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society 57 (1), 289–300.

Bojorges-Valdez, E., Yanez-Suarez, O., 2018. Association between EEG spectral power dynamics and event related potential amplitude on a P300 speller. Biomedical Physics and Engineering Express 4 (2).

Breitwieser, C., Pokorny, C., Müller-Putz, G. R., 2016. A hybrid three-class brain-computer interface system utilizing SSSEPs and transient ERPs. Journal of Neural Engineering 13 (6).

Costa, M., Goldberder, A. L., Peng, C.-K., 2002. Multiscale Entropy Analysis of Complex Physiologic Time Series. Physical Review Letters 89 (6), 1–4.

Costa, M., Goldberger, A. L., Peng, C. K., 2005. Multiscale entropy analysis of biological signals. Physical Review E - Statistical, Nonlinear, and Soft Matter Physics 71 (2), 1–18.

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

Goodale, M. A., Ingle, D. J., Mansfield, R. J. W., 1982. Analysis of visual behavior. Mit Press Cambridge, MA:.

He, S., Zhang, R., Wang, Q., Chen, Y., Yang, T., Feng, Z., Zhang, Y., Shao, M., Li, Y., 2017. A P300-Based Threshold-Free Brain Switch and Its Application in Wheelchair Control. IEEE Transactions on Neural Systems and Rehabilitation Engineering 25 (6), 715–725.

Huettel, S. A., McCarthy, G., 2004. What is odd in the oddball task? Prefrontal cortex is activated by dynamic changes in response strategy. Neuropsychologia 42 (3), 379–386.

Humeau-Heurtier, A., 2015. The multiscale entropy algorithm and its variants: A review. Entropy 17 (5), 3110–3123.

Jasper, H. H., 1958. The ten-twenty electrode system of the international federation. Electroencephalography and Clinical Neurophysiology 10, 371–375.

Lebedev, M. A., Messinger, A., Kralik, J. D., Wise, S. P., 2004. Representation of attended versus remembered locations in prefrontal cortex. PLoS Biology 2 (11).

Li, Y., Pan, J., Wang, F., Yu, Z., 2013. A hybrid BCI system combining P300 and SSVEP and its application to wheelchair control. IEEE Transactions on Biomedical Engineering 60 (11), 3156–3166.

Ma, Z., Qiu, T., 2018. Quasi-periodic fluctuation in Donchin's speller signals and its potential use for asynchronous control. Biomedizinische Technik 63 (2), 105–112.

Martínez-Cagigal, V., Gomez-Pilar, J., Álvarez, D., Hornero, R., 2017. An asynchronous P300-based Brain–Computer Interface web browser for severely disabled people. IEEE Transactions on Neural Systems and Rehabilitation Engineering 25 (8), 1332–1342.

Martínez-Cagigal, V., Santamaría-Vázquez, E., Gomez-Pilar, J., Hornero, R., 2019. Towards an accessible use of smartphone-based social networks through brain-computer interfaces. Expert Systems with Applications 120, 155–166.

Martínez-Cagigal, V., Santamaría Vázquez, E., Hornero, R., 2018. A Novel Hybrid Swarm Algorithm for P300-Based BCI Channel Selection. In: Proceedings of the World Congress on Medical Physics & Biomedical Engineering (Vol. 3) (IUPESM 2018). Springer, Prague, pp. 41–45.

Nicolas-Alonso, L. F., Gomez-Gil, J., 2012. Brain computer interfaces, a review. Sensors 12 (2), 1211–1279.

Panicker, R. C., Puthusserypady, S., Pryana, A. P., Sun, Y., 2010. Asynchronous P300 BCI: SSVEP-based control state detection. European Signal Processing Conference 58 (6), 934–938.

Pfurtscheller, G., 2010. The hybrid BCI. Frontiers in Neuroscience 4 (April), 1–11.
URL http://journal.frontiersin.org/article/10.3389/fnpro.2010.00003/abstract

Picton, T. W., 1992. The P300 wave of the human event-related potential. Journal of Clinical Neurophysiology 9 (4), 456–479.

Pinegger, A., Faller, J., Halder, S., Wriessnegger, S. C., Müller-Putz, G. R., 2015. Control or non-control state: that is the question! An asynchronous visual P300-based BCI approach. Journal of Neural Engineering 12 (1), 014001.

Richman, J. S., Moorman, J. R., 2000. Physiological time-series analysis using approximate entropy and sample entropy. American Journal of Physiology-Heart and Circulatory Physiology 278 (6), H2039–H2049.
URL http://www.physiology.org/doi/10.1152/ajpheart.2000.278.6.H2039

Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N., Wolpaw, J. R., 2004. BCI2000: A general-purpose brain-computer interface (BCI) system. IEEE Transactions on Biomedical Engineering 51 (6), 1034–1043.

Srinivasan, R., Bibi, F. A., Nunez, P. L., 2006. Steady-state visual evoked potentials: distributed local sources and wave-like dynamics are sensitive to flicker frequency. Brain Topography 18 (3), 167–187.
URL papers3://publication/uuid/FFD5DEE5-D729-4307-A600-6A24DCAA799F

Tang, J., Liu, Y., Jiang, J., Yu, Y., Hu, D., Zhou, Z., 2018. Toward Brain-Actuated Mobile Platform. International Journal of Human-Computer Interaction 00 (00), 1–12.
URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-85052079123{&}doi=10.1080{%}2F10447318.2018.1502000{&}partnerID=40{&}md5=e273e7a6b25ac3cc9f818703ae1191a3

Valencia, F., Porta, A., Clari, F., 2009. Refined Multiscale Entropy : Application to 24-h Holter Recordings of Heart Period Variability in Healthy and Aortic Stenosis Subjects. IEEE Transactions on Biomedical Engineering 56 (9), 2202–2213.

Witten, I. H., Frank, E., 2011. Data Mining: Practical Machine Learning Tools and Techniques, 3rd Edition. Morgan Kaufmann.

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., Vaughan, T. M., 2002. Brain-computer interfaces for communication and control. Clinical Neurophysiology 113 (6), 767–91.

Wolpaw, J. R., Wolpaw, E. W., 2012. Brain-computer interfaces: principles and practice. OUP USA.

Yu, Y., Zhou, Z., Jiang, J., Yin, E., Liu, K., Wang, J., Liu, Y., Hu, D., 2017a. Toward a Hybrid BCI: Self-Paced Operation of a P300-based Speller by Merging a Motor Imagery-Based "Brain Switch" into a P300 Spelling Approach. International Journal of Human-Computer Interaction 33 (8), 623–632.
URL http://dx.doi.org/10.1080/10447318.2016.1267450

Yu, Y., Zhou, Z., Liu, Y., Jiang, J., Yin, E., Zhang, N., Wang, Z., Liu, Y., Wu, X., Hu, D., 2017b. Self-paced operation of a wheelchair based on a hybrid brain-computer interface combining motor imagery and P300 potential. IEEE Transactions on Neural Systems and Rehabilitation Engineering 25 (12), 2516–2526.

Zhang, H., Guan, C., Wang, C., 2008. Asynchronous P300-based brain-computer interfaces: a computational approach with statistical models. IEEE Transactions on Biomedical Engineering 55 (6), 1754–63.
URL http://www.ncbi.nlm.nih.gov/pubmed/18714840

Brain–Computer Interface Channel Selection Optimization using Meta-heuristics and Evolutionary Algorithms

Víctor Martínez-Cagigal, Eduardo Santamaría-Vázquez, Roberto Hornero

Biomedical Engineering Group, E.T.S.I. Telecomunicación, University of Valladolid, Paseo de Belén 15, 47011, Valladolid, Spain

Applied Soft Computing
Volume -, -, Pages -, Under Review (R2)

Abstract

Many brain–computer interface (BCI) studies overlook the channel optimization procedure due to its inherent complexity. However, a suitable channel selection procedure increases the performance and users' comfort while reducing the cost of the system. Evolutionary meta-heuristics, which have gained importance due to their excellent performances in solving complex problems, have not been fully exploited yet in this context. The purpose of the study is two-fold: (1) to propose a novel algorithm to find an optimal channel set for each user and compare its usefulness with other existing meta-heuristics; and (2) to establish guidelines for adapting these optimization strategies to this framework. A total of 3 single-objective (GA, BDE, BPSO) and 3 multi-objective (NSGA-II, BMOPSO, SPEA2) existing algorithms have been adapted and tested with 3 public databases: 'BCI competition III–dataset II', 'Center Speller' and 'RSVP Speller'. Dual-Front Sorting Algorithm (DFGA), a novel multi-objective discrete method especially adapted to the BCI framework, is proposed as well. Results show that all the applied meta-heuristics reached accuracies that significantly outperformed the entire set and the common 8-channel set of P300-based BCIs. DFGA shown a significant improvement of accuracy of 3.9% over the 8-channel set using also 8 channels; and obtained similar accuracies than it using a mean of only 4.66 channels. Binary-based algorithms stood out because of their faster convergence, especially the DFGA. Topographic results shown that optimal sets differed among users, which reinforces the need to customize a channel set for each of them. The proposed method computes an optimal subset for each number of channels, allowing the user to select the most suitable set for further BCI sessions.

Keywords: Brain–computer interface (BCI), channel selection, multi-objective optimization, evolutionary algorithms, P300 event-related potentials.

1. Introduction

Brain–Computer Interfaces (BCIs) are communication systems that allow users to control devices and applications using their own brain signals. These systems have been successfully applied in order to improve the quality of life of motor-disabled people who suffer any disease that impairs the neural pathways that control muscles or even the muscles themselves (Wolpaw and Wolpaw, 2012). Electroencephalogram (EEG) is commonly used to monitor the brain activity due to its portability, non-invasiveness and low cost. Therefore, electric potentials are recorded by placing electrodes over the users' scalp (Wolpaw and Wolpaw, 2012).

Since the decoding of users' intentions from the EEG is not straightforward, many BCIs use exogenous signals to handle the control of the system. In particular, P300 evoked potentials, which are positive peaks produced as a response to infrequent and significant stimuli at about 300 ms after their onset, are the key aspect of the most known BCI-based spelling system (Wolpaw and Wolpaw, 2012). The 'P300 Speller' generates these signals through the odd-ball paradigm in order to spell certain words or commands. The application displays a matrix that contains characters or symbols, whose rows and columns are randomly flashing. Users, who have to focus attention on a desired command, will generate a P300 potential whenever the row or the column that contains it is highlighted. Hence, the selected command is determined by computing the intersection between the row and the column that produced the potential (Farwell and Donchin, 1988).

Due to the low signal-to-noise ratio and the high inter-session variability of these event-related potentials, several repetitions of the same stimulus are required to detect a reliable potential. Without a proper processing stage, this high dimensional data may produce an overfitting of the classifier, resulting in a reduced performance (Cecotti et al., 2011; Perseh and Sharafat, 2012). The curse of dimensionality can be addressed by means

Email addresses: victor.martinez@gib.tel.uva.es (Víctor Martínez-Cagigal), eduardo.santamaria@gib.tel.uva.es (Eduardo Santamaría-Vázquez), robhor@tel.uva.es (Roberto Hornero)

DOI: Under Review (R2)

of feature selection and extraction methods (Perseh and Sharafat, 2012; Tahernezhad-Javazm et al., 2017), regularized classifiers (Blankertz et al., 2011) or channel selection procedures (Cecotti et al., 2011; Martínez-Cagigal and Hornero, 2017b). Among them, only channel selection methods are able to reduce the cost of the system, reduce the power consumption in EEG caps and increase the users' comfort (Cecotti et al., 2011). Nevertheless, the selection of the most relevant sensors is not trivial, since there are 2^N subset combinations for an N-channel cap, making the exhaustive search intractable in practice (Cecotti et al., 2011). For this reason, most P300-based studies overlook the optimization of the most relevant channel subset and take a predefined 8-channel set as a general rule of thumb (Krusienski et al., 2008). Notwithstanding its usefulness as a quick solution, an optimization for each user is beneficial owing to the intrinsic inter-subject variability of the BCI systems.

Although there are well-known feature selection methods that could be applied to this problem, such as stepwise regression (Jobson, 1991), fast correlation based filters (Yu and Liu, 2003) or even elastic neural networks (Zou and Hastie, 2005), meta-heuristics have demonstrated high performances solving complex optimization problems (Yang, 2014). Heuristics refer to problem-specific strategies that iteratively improve a candidate solution, whereas meta-heuristics generalize these strategies to problem-independent frameworks (Yang, 2014; Bozorg-Haddad et al., 2017). Swarm intelligence techniques and evolutionary algorithms, families of population-based meta-heuristics, have been previously applied to channel selection optimization both in motor imagery (MI) (Lv and Liu, 2008; Hasan and Gan, 2009; Hasan et al., 2010; Wei and Wang, 2011; Kee et al., 2015; Aler and Galván, 2015; Franklin Alex Joseph and Govindaraju, 2019; Zhang and Wei, 2019; González et al., 2019) and P300-based BCIs (Perseh and Sharafat, 2012; Jin et al., 2010; Gonzalez et al., 2013, 2014; Kee et al., 2015; Martínez-Cagigal and Hornero, 2017b; Chaurasiya et al., 2017; Arican and Polat, 2020). Regarding the P300-based BCI studies, most of them have used single-objective algorithms that optimized the final classification accuracy of the system (Jin et al., 2010; Perseh and Sharafat, 2012; Gonzalez et al., 2013, 2014; Arican and Polat, 2020). However, a channel selection procedure should follow a two-fold objective: (i) to minimize the number of selected channels, and (ii) to maximize the system's performance. Some recent studies used a weighted aggregation approach to combine both objectives into a single one, but the simultaneous optimization was not explored (Martínez-Cagigal and Hornero, 2017b; Martínez-Cagigal et al., 2018; Zhang and Wei, 2019).

Traditional multi-objective approaches, which optimize both objectives at the same time, have been explored in MI-based BCIs, such as multi-objective particle swarm optimization (MOPSO) (Hasan and Gan, 2009; Hasan et al., 2010; Wei and Wang, 2011) or non-sorting genetic algorithm II (NSGA-II) (Aler and Galván, 2015; González et al., 2019). However, MI-based and P300-based BCIs are completely different in terms of processing and signal elicitation and thus, methodology and results cannot be generalized. By contrast, multi-objective algorithms in P300-based BCIs are more limited. Kee et al. (2015) compared the performance between several single-objective genetic algorithms (GA) and NSGA-II with 2 subjects, whereas Chaurasiya et al. (2017) employed a multi-objective binary differential-evolution algorithm with 9 subjects, reaching several subsets of channels that assured suitable classification performances. Nevertheless, the number of subjects was limited, and both databases were recorded using the row-col paradigm (RCP). Nowadays, P300-based BCIs offer a wide range of stimulation paradigms that elicit different event-related responses and thus, the generalization of those results to other setups is limited. Furthermore, despite their scarce application in P300-based BCI studies, swarm intelligence and evolutionary computation are growing research fields that integrate a large amount of different algorithms that could be adapted to the channel selection problem. In fact, the vast majority of them have not yet been applied in P300-based BCIs. To the best of our knowledge, there are not studies that compare their efficacy to select the most appropriate channel subset or even establish the key aspects for their adaptation to BCI systems, which is not trivial. Moreover, there is also no study aimed at designing any multi-objective algorithm customized for the P300-based BCI channel selection problem.

The objective of this study is two-fold: (1) to propose a novel multi-objective method to find an optimal channel set especially suited for P300-based BCIs and compare its usefulness with 6 additional meta-heuristics; and (2) to establish guidelines for adapting these optimization strategies to the channel selection problem. Although there are many meta-heuristics that could be adapted to this problem, only those that have direct or explicit contribution to our proposed multi-objective meta-heuristic were included in this comparison: GA, BDE and BPSO as single-objective; and NSGA-II, SPEA2 and BMOPSO as multi-objective. We have also tried to maintain diversity in the way they deal to the updating of the population for each iteration. The manuscript is organized as follows. Section 2 details the 3 different databases that were used to test the algorithms. In section 3, the methodology is described, including pre-processing and feature extraction, problem definition, 3 single-objective and 3 multi-objective meta-heuristics, and our proposed algorithm. Sections 4 and 5 show the results and discuss them, respectively; analyzing the convergence, selected channel distributions, assessment accuracies, hyperparameters and computational cost. At the end of section 5, a set of guidelines for adapting meta-heuristics to the BCI channel selection problem is detailed, followed by the contributions, limitations and future endeavors. To sum

up, section 6 draws the final conclusions of the study.

2. Subjects

In order to improve the generalization of the results, algorithms have been tested with three public P300-based BCI databases that used different stimulation paradigms: row-col paradigm (RCP), center speller (CS) and rapid serial visual presentation (RSVP). Examples of stimulation sequences for the paradigms are depicted in the figure 1.

2.1. BCI competition III: dataset II

The 'BCI competition III: dataset II' (Blankertz et al., 2006) was recorded from 2 different healthy subjects (i.e., A and B) that were asked to spell words in 5 RCP sessions. Signals were recorded using a 64-channel EEG cap with a sampling frequency of 240 Hz and band-pass filtered from 0.1 Hz to 60 Hz. Training and testing sets were composed of 85 and 100 trials, respectively (Blankertz et al., 2006). RCP is the most common P300-based spelling paradigm, which consists of displaying a matrix that contains characters or symbols. Users have to stare at the target command while the matrix's rows and columns are randomly flashing. Whenever the row or column that contains the target is flickered, a P300 potential is generated. Hence, the desired command can be identified by computing the intersection between the row and the column that produced these P300 responses (Farwell and Donchin, 1988). In this dataset, there are 12 different classes (i.e., rows and columns), and 15 sequences (i.e., repetitions) were used. Therefore, a trial is composed by 180 observations (Blankertz et al., 2006).

2.2. Center Speller database

The 'Center Speller (008-2015)' database (Treder et al., 2011) was recorded from 13 healthy subjects (i.e., C01-C13) that were asked to perform spelling tasks using the CS paradigm. Signals were recorded using a 63-channel EEG cap with a sampling frequency of 250 Hz and band-pass filtered from 0.016 Hz to 250 Hz. Training data was composed of 17 trials, whereas testing data varied between 32-49 trials, depending on the subject (Treder et al., 2011). CS was originally designed to avoid eye movements. The paradigm displays groups of commands in the center of the screen, superimposed with colored geometric shapes. The groups are randomly flickered until the user selects one of them. Then, the commands that were included inside the selected group are displayed in the same way, allowing the user to select the final command (Treder et al., 2011). In practice, there are 12 different classes (6 groups in 2 levels), and 10 sequences were used. A trial is composed by 120 observations (Treder et al., 2011).

2.3. RSVP Speller database

The 'RSVP Speller (010-2015)' database (Acqualagna and Blankertz, 2013) was recorded from 12 healthy subjects (i.e., R01-R12) that were asked to perform spelling tasks using the RSVP paradigm. Signals were recorded using a 63-channel EEG cap with a sampling frequency of 1000 Hz, and then down-sampled to 200 Hz (Acqualagna and Blankertz, 2013). However, since the fifth subject only used 61 channels, electrodes P8 and O2 were excluded from the database for the sake of homogeneity. Training data was composed of 24 trials, whereas testing data (copy and free spelling) varied between 37-50 trials, depending on the subject (Acqualagna and Blankertz, 2013). RSVP was also developed to exploit the foveal visual field and avoid eye movements by depicting symbols in the center of the screen in a serial manner. The database encompasses a vocabulary of 30 characters (26 letters and 4 symbols). In order to favor the identification of the shapes, half of the letters were uppercase and the other half lowercase, using 5 different colors. Therefore, there are 30 classes, and 10 sequences were used, resulting in 300 observations per trial (Acqualagna and Blankertz, 2013).

3. Methods

3.1. Pre-processing and feature extraction

Before applying any optimization procedure, relevant features of the EEG signals should be extracted for each epoch (i.e., stimulus) and channel. In fact, preprocessing, as well as feature extraction and selection procedures influence the final accuracy in a high extent. Due to the purpose of the study, signal processing stages were composed of a standard framework, intended to isolate the channel selection procedure. We did not applied any further pre-processing step besides the aforementioned band-pass filtering embedded in each database (Blankertz et al., 2006; Treder et al., 2011; Acqualagna and Blankertz, 2013). Epochs were extracted using windows in the range $[0, 800]$ ms from the stimuli onsets, and normalized using a z-score over a $[-200, 0]$ ms baseline (i.e., zero mean, unit standard deviation). As stated in BCI literature, this range is large enough to capture the relevant event-related potentials, including the P300 wave (Wolpaw and Wolpaw, 2012). These epochs were then decimated to 25 Hz, keeping a total of 20 features per stimulus and channel. It is noteworthy that the decimation process encompasses a low-pass filtering (to avoid aliasing), followed by a down-sampling procedure (Martínez-Cagigal and Hornero, 2017a,b). Hereafter, epochs from different databases and sampling rates have the same number of features. Note that, from the point of view of a subsequent classifier, epochs are input observations.

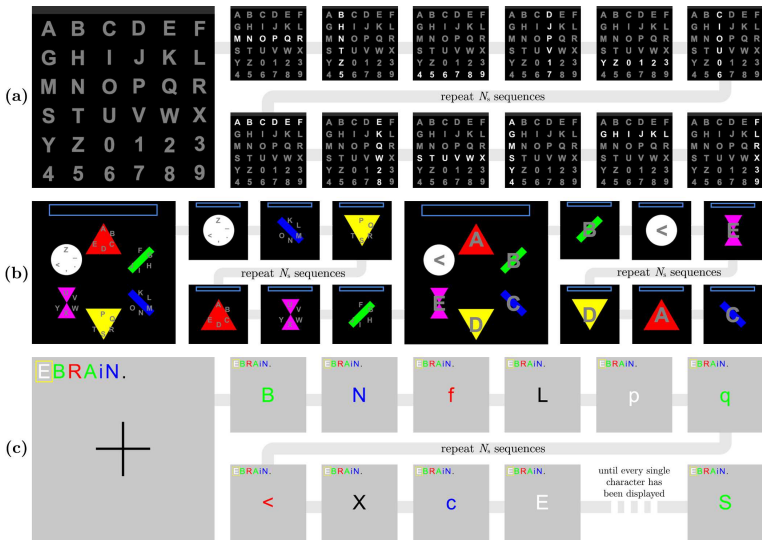


Figure 1: Examples of stimuli intensification sequences for the paradigms: (a) row-col paradigm, (b) center speller, (c) rapid serial visual presentation.

3.2. Defining the optimization problem

The goal of an optimization algorithm is to provide a suitable solution that satisfies the problem constraints and optimizes (either maximizing or minimizing) one or more objective functions to the greatest extent (Bozorg-Haddad et al., 2017). Since we are considering an N-channel selection problem, a possible solution may be defined as $\boldsymbol{x} = [x_1, x_2, \ldots, x_N], x_i \in \{0, 1\}$, where 1 and 0 represent the selection and rejection of a specific channel, respectively. Hence, the problem is constrained to a discrete N-dimensional space, whose solutions are restricted to binary positions. When a solution $\boldsymbol{x}$ is evaluated, the features associated with the channels that satisfy $x_i = 1$ are concatenated as an input feature vector.

In a BCI channel selection problem, two main objectives must be pursued: (i) to maximize the system's performance, and (ii) to minimize the number of channels. Even though the modeling of the latter is straightforward (see equation 1), the system's performance can be estimated following several approaches. The most intuitive solution is to use the output training accuracy of the classifier using a certain solution $\boldsymbol{x}$ (Jin et al., 2010; Kee et al., 2015; Chaurasiya et al., 2017). However, due to the limited number of trials, this method usually provides a low-resolution score (Colwell et al., 2014). The resolution can be improved using methods based on the stimuli, instead of the character trials. Previous studies have employed several approaches derived from the confusion matrix of the stimuli classification (Perseh and Sharafat, 2012; Gonzalez et al., 2013, 2014). Nevertheless, the area under ROC curve (AUC) is recommended because it is able to successfully estimate the discriminative ability of a binary classifier using only training data (Cecotti et al., 2011; Colwell et al., 2014). Therefore, the objectives may be modeled as follows:

$$\min F(\boldsymbol{x}) = \begin{cases} f_1(\boldsymbol{x}) = 1 - \overline{AUC}(\boldsymbol{x}) \\ f_2(\boldsymbol{x}) = \sum_{n=1}^{N} x_n \end{cases}, \qquad (1)$$

where $f_1(\boldsymbol{x})$ belongs to the first objective (i.e., minimize the error of the system) and $f_2(\boldsymbol{x})$ to the second objective (i.e., minimize the number of channels). In this study, $\overline{AUC}$ has been derived from a 5-fold cross-validated linear discriminant analysis (LDA) that is applied to the solution $\boldsymbol{x}$ using the training dataset (Martínez-Cagigal and Hornero, 2017a,b). That is, the features whose channels satisfy $x_i = 0$ are removed from the observations matrix, which is the input of the LDA classifier. Training set is then divided into 5 subsets and a cross-validation procedure is applied (i.e., 4 subsets are used for training and the remaining one for testing), returning a total of 5 AUCs. Finally, $\overline{AUC}$ is computed as the average of all of them. LDA was used as classifier due to its well-known excellent performances in P300-based BCI and the lack of hyperparameters to optimize (Tahernezhad-Javazm et al.,

2017; Blankertz et al., 2011; Kee et al., 2015; Aler and Galván, 2015; Chaurasiya et al., 2017; Treder et al., 2011; Martínez-Cagigal and Hornero, 2017b; Martínez-Cagigal et al., 2017).

3.3. Single-objective meta-heuristics

Meta-heuristics produce acceptable solutions to complex problems in a reasonable computation time (Yang, 2014). In particular, single-objective meta-heuristics iteratively produce these solutions following a certain objective. However, a BCI channel selection problem should have a two-fold purpose. Thus, the multi-objective problem stated in equation (1) is then combined into a single-objective one (Coello and Reyes-Sierra, 2006):

$$\min F(\boldsymbol{x}) = \omega_1 f_1(\boldsymbol{x}) + \omega_2 \left(\frac{f_2(\boldsymbol{x}) - 1}{N - 1} \right)^3, \qquad (2)$$

where $\omega_1 + \omega_2 = 1$, and ω_1 and ω_2 are constants that weigh the importance of each objective. Since we consider that reaching suitable accuracies is more important than drastically reducing the number of required channels, coefficients have been heuristically set to $\omega_1 = 0.7$ and $\omega_2 = 0.3$ (Martínez-Cagigal and Hornero, 2017a,b; Martínez-Cagigal et al., 2018). In addition, after mapping the $f_2(\boldsymbol{x})$ from $[1, N] \rightarrow [0, 1]$, its output is raised to the 3-th power in order to empathize the search of lightweight solutions. Note that the polynomial function punishes more the search of solutions with high number of channels than a simple linear function. This function was heuristically chosen after a preliminary testing (Martínez-Cagigal and Hornero, 2017a,b; Martínez-Cagigal et al., 2018). The three single-objective meta-heuristics that have been adapted to BCI framework are described below.

3.3.1. Genetic Algorithm

One of the most well-known meta-heuristics is the genetic algorithm (GA), originally developed by Holland (1992). GAs have been modified to improve their ability to find the global optimum of a complex optimization problem in many ways. GAs apply the Darwinian principle of survival of the fittest individuals of a population using recombination, selection and mutation operators (Yang, 2014; Bozorg-Haddad et al., 2017). In this study, a GA with elitism, binary tournament selection, single-point crossover and bit string mutation has been employed (Yang, 2014; Bozorg-Haddad et al., 2017).

3.3.2. Binary Differential Evolution

The differential evolution (DE) algorithm, originally developed by Storn and Price (1997) for continuous functions, has some similarities to GAs owing to its structure, composed by mutation, crossover and selection. However, instead of making random mutation and crossover schemes, DE combines the information of three randomly chosen individuals. Binary DE (BDE) applies a discretization of the mutation formula in order to adapt it to binary problems (Wang et al., 2012). The mutation of the i-th channel of an individual $\boldsymbol{x}$ is performed as follows:

$$x_i' = \begin{cases} u_i, & \text{if rand} \leq p_c \text{ or } i = r \\ x_i, & \text{otherwise} \end{cases}, \qquad (3)$$

where rand $\sim \mathcal{U}(0, 1)$, r is a random integer between $[1, N]$, p_c is the crossover rate, and u_i is the mutated channel, computed as:

$$u_i = \begin{cases} 1, & \text{if rand} \leq (1 + e^{\frac{-2b(v_i + F \cdot (y_i - z_i) - 1/2)}{1+2F}})^{-1} \\ 0, & \text{otherwise} \end{cases}, \qquad (4)$$

where rand $\sim \mathcal{U}(0, 1)$; $\boldsymbol{v}$, $\boldsymbol{y}$ and $\boldsymbol{z}$ are randomly selected individuals of the current population; F is the weighting factor; and $b > 0$ is the bandwidth factor.

3.3.3. Binary Particle Swarm Optimization

Kennedy and Eberhart (1995) developed the Particle Swarm Optimization (PSO) algorithm, a nature-inspired meta-heuristic based on the social schooling and flocking behavior of fishes and birds. The optimization is based on adjusting the trajectories and positions of a set of particles (i.e., solutions) that "fly" over the search space, whose movement have both deterministic and stochastic components (Kennedy and Eberhart, 1995; Yang, 2014; Bozorg-Haddad et al., 2017). In this study, the standard constraint of Clerc and Kennedy (2002) is used, which leads to:

$$\boldsymbol{v}' = \chi[\boldsymbol{v} + \epsilon_1 C_1(\boldsymbol{l} - \boldsymbol{x}) + \epsilon_2 C_2(\boldsymbol{g} - \boldsymbol{x})], \qquad (5)$$

$$\chi = \frac{2}{\phi - 2 + \sqrt{\phi^2 - 4\phi}}, \text{ with } \phi = C_1 + C_2; \qquad (6)$$

where $\boldsymbol{v}'$ is the updated velocity of a particle $\boldsymbol{x}$; $\boldsymbol{v}$ is the last velocity; $\epsilon_1, \epsilon_2 \sim \mathcal{U}(0, 1)$; χ is the constraint multiplier; C_1 and C_2 are the personal and global confidence constants, respectively; $\boldsymbol{l}$ is the best position found by the particle; and $\boldsymbol{g}$ is the best global position found so far. It is worthy to note that the standard constraint requires that $\phi > 4$ (Clerc and Kennedy, 2002; Poli et al., 2007). Since the velocities are continuous, the algorithm should be adapted to binary spaces. Binary PSO (BSPO) is usually achieved using a position transformation via transfer functions (Kennedy and Eberhart, 1997; Mirjalili and Lewis, 2013). In this study, the adaptation has been performed following the expression:

$$x_i' = \begin{cases} \neg x_i, & \text{if rand} < T(v_i') \\ x_i, & \text{if rand} \geq T(v_i') \end{cases}, \qquad (7)$$

where rand $\sim \mathcal{U}(0, 1)$, and $T(t) = |t/\sqrt{1 + t^2}|$ is a v-shaped transfer function (Mirjalili and Lewis, 2013).

3.4. Multi-objective meta-heuristics

In contrast to the single-objective strategies, multi-objective meta-heuristics involve the simultaneous optimization of more than one objective. Since these objectives are usually conflicting among themselves, the concept of dominance is introduced for determining the quality of each solution (Deb, 2005). It is said that a solution $\boldsymbol{y}$ dominates a solution $\boldsymbol{z}$ (i.e., $\boldsymbol{y} \succ \boldsymbol{z}$) if $\forall i : f_i(\boldsymbol{y}) \leq f_i(\boldsymbol{z})$ and $\exists j : f_j(\boldsymbol{y}) < f_j(\boldsymbol{z})$. The Pareto-front, a curve that contains optimal solutions (i.e., those that are not dominated by any other solutions), is estimated by the multi-objective algorithms, depicting the trade-off among the objectives (Deb, 2005). Regarding the BCI channel selection problem, the Pareto-front returns a set of solutions that have different number of required channels, allowing the user to select one of them.

3.4.1. Non-Sorting Genetic Algorithm II

The most popular approach for extending GAs to multi-objective optimization problems is the Non-Sorting Genetic Algorithm II (NSGA-II), proposed by Deb et al. (2002). Crossover and mutation operators are the same as GAs, whereas the selection operator is more complex. Firstly, in order to estimate the quality of each chromosome, the algorithm establishes a hierarchy of Pareto-fronts according to its dominance. The first Pareto-front (i.e., rank $= 1$) is composed by the non-dominated chromosomes of the current population. Then, the second Pareto-front (i.e., rank $= 2$) is computed in the same way, but ignoring the chromosomes of the first front. This process is sequentially repeated until there are no chromosomes left (Deb et al., 2002). However, the selection of a parent population is not only based on the rank of the chromosomes, but also on their crowding distances. These metrics are included in order to spread the solutions along the Pareto-front and avoid getting trapped in local minima. The crowding distance of a chromosome is computed as the average distance between its two adjacent solutions with the same rank. Boundary solutions are assigned an infinite distance value. Considering two chromosomes, the solution with lower rank is preferred. Whether both have the same rank, the less crowded solution is preferred (i.e., higher distance value). The parent population is sequentially filled with the firsts Pareto-fronts until the number of included solutions is greater or equal than $m/2$. Then, parent solutions are truncated based on the crowding distances until the number of solutions is exactly $m/2$. Further information can be found in Deb et al. (2002).

3.4.2. Binary Multi-Objective PSO

Due to its usefulness to solve complex optimization problems, many authors have tried to adapt the PSO algorithm to multi-objective environments (Coello and Reyes-Sierra, 2006). Here, a Binary Multi-Objective PSO (BMOPSO) approach is applied. Since the conflicting objectives do not permit the establishing of an optimal global solution $\boldsymbol{g}$, the major adaptation should reside in the way to select the leader of each particle. In this study, a repository approach is employed. Non-dominated solutions are stored in an external repository with "unlimited" size. Note that its maximum size would be the maximum number of channels (i.e., the resolution of the BCI problem). A particle's leader is randomly selected from the repository, and it is attached to the particle until the leader is no longer part of the repository. In that case, the leader is substituted by another randomly selected one. In addition, a three-fold bit string mutation is also used, which consists on dividing the swarm in three parts and apply: (1) no mutation; (2) uniform mutation with probability p_m; (3) non-uniform mutation with probability $p_n = (1 - \text{gen}/\text{ngen})^{5N}$ (Reyes-Sierra and Coello, 2005).

3.4.3. Strength Pareto Evolutionary Algorithm 2

Zitzler et al. (2001) developed a multi-objective algorithm that integrates the concepts of dominance and crowding density in a single metric, called Strength Pareto Evolutionary Algorithm 2 (SPEA2). In order to compute the unified metric, the strength concept is introduced. The strength S_i is computed as the number of solutions that the i-th particle dominates. Then, the unified fitness is calculated as follows:

$$F_i = R_i + \frac{1}{\sigma_i^k + 2}, \tag{8}$$

where R_i is the sum of the strengths of the particles that dominates i, and σ_i^k is the distance sought of the particle (i.e., distance to the k-nearest neighbor), where $k = \lfloor\sqrt{m}\rfloor$. Note that non-dominated individuals would have $R = 0$ and thus, $F < 1$. SPEA2 also uses a repository with fixed size that is updated following an environmental selection procedure. Solutions are sorted according to their F values, and the repository is filled with them. If the number of solutions of the repository is higher than its maximum size N_r, a truncation process is applied. Then, the algorithm removes solutions from the repository according to their σ^k (i.e., high σ_k values are preferred), in order to preserve Pareto-front spreading (Zitzler et al., 2001).

3.5. Our proposal: Dual-Front Genetic Algorithm

Even though there is a large variety of meta-heuristics from single to multi-objective algorithms, all of them should be adapted to the channel selection problem. The BCI framework forces the algorithms to work with binary solutions, involving the use of transfer functions in some cases. These functions convert a defined alteration of a solution into a probability of change, increasing the stochasticity of the algorithm. Moreover, the conversion can be addressed as a multivalued function of the type f :

$\mathbb{R} \rightarrow \{0, 1\}$, which means that there are infinite input values that produce exactly the same output, hindering the local exploitation of new solutions. By extension, there is no point in using operators based on continuous distances. Furthermore, since $f_2(\boldsymbol{x})$ already restricts the size of multi-objective repositories to N, limitation strategies (e.g., crowding, distance sought) also entail an unnecessary computational cost. In order to overcome these restraints, a novel multi-objective algorithm is proposed: the Dual-Front Genetic Algorithm (DFGA). DFGA is specially designed to the BCI framework by means of five key aspects: (i) deterministic initialization, (ii) elitism, (iii) dual-front sorting, (iv) genetic operators, and (v) synthetic solutions. DFGA pseudocode is depicted in the figure 2(f).

Deterministic initialization. Heuristics generally initialize the population by generating random solutions. However, the use of deterministic initialization may reduce both the inter-run variability due to stochastic effects and a large amount of computation time. Although deterministic algorithms are not likely to provide a global optimum, DFGA considers their outputs as intermediate solutions. Regardless of their qualities, we hypothesize that these solutions are equivalent to those that will be eventually reached after several generations of a randomly-initialized algorithm. In this study, backward elimination (BE) is used to initialize the repository. The algorithm begins with the full set of channels and sequentially removes the most irrelevant one (Jobson, 1991). The rejected channel in each step is the one that returns the minimum $f_1(\boldsymbol{x})$ value if removed from the model $\boldsymbol{x}$ (i.e., its inclusion does not contribute to improve the system's performance). The algorithm continues removing channels until the set is empty. Note that this operation will fill the repository $\mathcal{R}$ up with N solutions.

Elitism. In each generation, the repository is updated following an elitist approach. As depicted in the figure 2(a), for each unique value of $f_2(\boldsymbol{x})$ (i.e., for each number of channels), the repository solution that minimizes $f_1(\boldsymbol{x})$ is selected. Note that this operation is applied in the repository, which includes both non-dominated and dominated solutions, creating a balance between local and global exploitation.

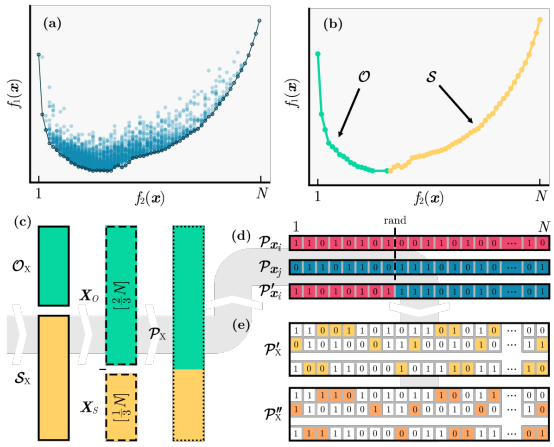
Dual-front sorting. Due to the deterministic initialization, the repository should have a well-defined curve from the very beginning of the algorithm. This aspect leads to a Pareto-front that is supposed to include solutions with few number of channels. Traditionally, only the Pareto-optimal solutions are considered in the selection stage. Despite their convenience over dominated solutions, considering only the Pareto-front will lead to a local exploitation of solutions with few channels. However, because of the intrinsic fixed size of the repository in BCI problems (i.e., limited to N), the exploitation of solutions with a greater number of channels is no longer an issue. Furthermore, it may favor the spreading of the Pareto-front and the global search of DFGA. According to this rationale, DFGA subdivides the repository into two sets: $\mathcal{O}$ (i.e., optimal set), which includes the non-dominated solutions; and $\mathcal{S}$ (i.e., sub-optimal set), which includes the dominated solutions. Dual-front sorting operation is shown in figure 2(b). Then, binary tournament selection is applied in both sets, selecting $^{2N}/_{3}$ solutions from $\mathcal{O}$, and $^{N}/_{3}$ solutions from $\mathcal{S}$. Note that a solution may be selected more than once in the new population. Finally, these solutions are combined in the population in order to suffer recombination (i.e., crossover) and mutation, as shown in the figure 2(c).

Genetic operators. Owing to the binary nature of the search space, we consider that traditional genetic operators are the most convenient approach for generating new solutions from a parent population. First, for each solution $\boldsymbol{x}_i$, single-point crossover is applied with probability p_c. That is, $\boldsymbol{x}_i$ and another randomly picked solution $\boldsymbol{x}_j$ ($i \neq j$) are combined into $\boldsymbol{x}'_i \leftarrow \boldsymbol{x}_i[1:u] \cup \boldsymbol{x}_j[u+1:N]$, where $u \sim \text{rand} \in [1, N]$. For each solution, bit-string mutation is also computed with probability p_m. In other words, if the n-th bit of a solution $\boldsymbol{x}'_i$ has to be mutated, its value is flipped (i.e., $\boldsymbol{x}''_i[n] \leftarrow \neg \boldsymbol{x}'_i[n]$). The procedure is illustrated in the figure 2(d–e).

Synthetic solutions. When the values of p_c or p_m are too high, the mutated population tends to exploit the middle part of the repository. In other words, solutions with few channels tend to add more channels, whereas crowded solutions tend to decrease their number of channels. In order to maintain a similar exploitation across the entire repository spectrum, synthetic solutions are generated apart from the mutated population. However, a random generation of solutions across this spectrum will unnecessarily increase the number of evaluations, slowing down the algorithm. DFGA generates synthetic solutions using a metric that is intended to maintain the most relevant channels of the current repository. The rank of the i-th channel is defined as the number of times that the channel i is present in the repository (i.e., $r_i = |i \in \mathcal{R}|$). DFGA iteratively creates solutions that have from 1 to $N-1$ channels by means of a roulette wheel selection (i.e., fitness proportionate selection) based on the rank values. It is worthy to mention that DFGA generates a total of $N-1$ solutions, since the N-th solution that contains all the channels is already part of the repository.

4. Results

For comparison purposes, the number of individuals of every single meta-heuristic was fixed to $m = 20$. Hyperparameters, detailed in table 1, were set following the recommendations of the literature (Deb et al., 2002; Wang et al., 2012; Clerc and Kennedy, 2002). In order to assure



(f) DFGA algorithm

```
procedure DFGA(N, p_c, p_m)
    [R_{X,F}] ← backward_elimination()
    [O_{X,F}, S_{X,F}] ← dual_front(R)
    while ¬stopCondition do
        /* Parent selection */
        X_o ← binary_tournament(2N/3, O)
        X_s ← binary_tournament(N/3, S)
        P_X ← X_o ∪ X_s
        /* Genetic operators */
        P'_X ← crossover(P_X, p_c)
        P''_X ← mutation(P'_X, p_m)
        P''_F ← F(P''_X)
        /* Synthetic solutions */
        Q_X ← generate_synthetic(R_X)
        Q_F ← F(Q_X)
        /* Repository updating */
        R_{X,F} ← rep_update(R ∪ P'' ∪ Q)
        [O_{X,F}, S_{X,F}] ← dual_front(R)
    end while
    return O_{X,F}
end procedure
```

Figure 2: Visual aids to clarify DFGA operations: (a) elitist repository updating, (b) dual-front sorting, (c) parent selection, (d) single-point crossover, and (e) bit-string mutation. The pseudocode of the algorithm is depicted in (f).

Table 1: Method-specific hyperparameters.

Prm.	Value	Description	Algorithm
m	20	No. individuals	All
p_m	$1/N$	[a]Mutation rate	GA, NSGA-II, BMOPSO, SPEA2, DFGA
p_c	0.90	[a]Crossover rate	GA, NSGA-II, SPEA2, DFGA
F	0.80	[b]Weighting factor	BDE
b	N	[b]Bandwidth factor	BDE
p_{de}	0.20	[b]BDE crossover rate	BDE
C_1	2.05	[c]Personal confidence	BPSO, BMOPSO
C_2	2.05	[c]Global confidence	BPSO, BMOPSO
V_{max}	1.00	[c]Maximum velocity	BPSO, BMOPSO

[a]Deb et al. (2002), [b]Wang et al. (2012), [c]Clerc and Kennedy (2002).

Table 2: Approximate computational costs of single and multi-objective meta-heuristics.

	Mtd.	No. eval.	Eval. time	No. gen.
Single	GA	20 eval./gen.	785 ms/eval.	200 gen.
	BDE	20 eval./gen.	810 ms/eval.	200 gen.
	BPSO	20 eval./gen.	858 ms/eval.	200 gen.
Multi	NSGA-II	40 eval./gen.	331 ms/eval.	100 gen.
	SPEA2	20 eval./gen.	835 ms/eval.	200 gen.
	BMOPSO	20 eval./gen.	852 ms/eval.	200 gen.
	DFGA	123 eval./gen.	591 ms/eval.	32 gen.

Mtd.: method, gen.: generation, eval.: evaluation.

a fair comparison among the algorithms, the number of generations varied in function of the amount of evaluations that were performed in a single iteration. Table 2 details the computational cost, including the amount of evaluations per generation and the number of generations for each method. In total, 4000 evaluations were performed. Furthermore, all the algorithms were computed 20 times in order to avoid local minima. The experiments were executed in an Intel Core i7-7700 CPU @ 3.60 GHz, 32GB RAM, Windows 10 Pro, using MATLAB® 2018b.

A convergence analysis for single-objective meta-heuristics is depicted in the figure 3. These averaged convergence curves show the evolution of the aggregated objective function $F(\boldsymbol{x})$ across the generations. Thus, they estimate the ability of each method to search for an optimal solution in the training phase. The detailed convergence curves for each subject can be found in the supplementary material. Concerning the multi-objective meta-heuristics, the evolution of the computed Pareto-fronts over the generations of the algorithms is depicted in the figure 4, also in training phase.

Ranks of selected channels for both single and multi-objective meta-heuristics are displayed in the figure 5, including the common Krusienski's 8-channel set. The rank of a channel is defined as the normalized number of times that the channel was selected in the repetitions. For multi-objective algorithms, only the ranks of channels that belongs to the repository are included. Scalp distributions of the averaged rank values over the meta-heuristics are depicted for each subject as well.

In order to evaluate the real performance of the single-objective algorithms using testing datasets, it is required

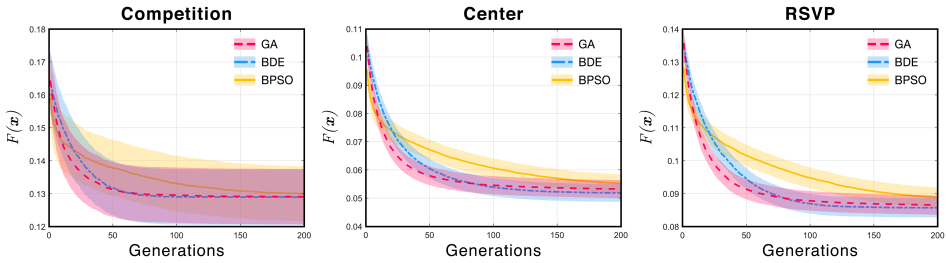


Figure 3: Averaged convergence curves of single-objective meta-heuristics (GA, BDE and BPSO) for each database in function of the $F(\boldsymbol{x})$ aggregated function. Mean values are displayed with solid lines, whereas the 95% confidence interval of the subjects' repetitions is indicated by the shaded area.

Table 3: Averaged testing accuracies and number of channels across users of the selected run for each single-objective method.

Mtd.	Competition		Center		RSVP	
	Acc.	N	Acc.	N	Acc.	N
GA	92.0%	14.0	97.4%	12.4	84.6%	13.4
BDE	92.0%	14.5	97.9%	12.5	85.5%	13.4
BPSO	92.0%	14.0	96.8%	12.5	85.0%	13.7
ALL	92.0%	64.0	86.5%	63.0	80.3%	61.0
KRU	86.5%	8.0	95.2%	8.0	78.6%	8.0

Mtd.: method, Acc.: accuracy, N: no. of sequences. Results obtained using the maximum number of sequences available for each database (competition: 15, center: 10, RSVP: 10).

to select a single solution among the repetitions. Therefore, the solution that reached the minimal $F(\boldsymbol{x})$ value was selected for each single-objective method. Table 3 summarizes the averaged testing accuracies and number of channels of the selected solutions for each subject, in function of the employed method, using the maximum number of sequences available in each database. Regarding the multi-objective algorithms, once the repetitions are computed, the final Pareto-front for each subject is composed by the non-dominated solutions among the returned in each repetition. Testing accuracies (i.e., ratio of correctly predicted characters) of the solutions that belongs to the final Pareto-fronts are shown in the figure 6, again using the maximum number of sequences available. Finally, computation costs of all algorithms are detailed in the table 2.

5. Discussion

5.1. Convergence analysis

Regarding the single-objective meta-heuristics, results show that the inherently discrete algorithms (i.e., GA and BDE) converge to optimal solutions faster than BPSO, and are able to reach the minimal objective value for every single subject. Inherent discrete algorithms are understood as meta-heuristics that employs binary methodologies to improve their solutions (i.e., mutation, crossover). Even though BPSO shows a slower convergence than GA or BDE, the reached $F(\boldsymbol{x})$ values are almost analogous, suggesting that BPSO, GA and BDE will show similar performances in testing phase. It is also noteworthy that, even though the averaged convergence of GA is faster than BDE, the curve eventually reach a standstill over the 100th generation, being overcome by BDE from than point on.

Multi-objective meta-heuristics results show that DFGA, NSGA-II and SPEA2 algorithms are able to reach similar Pareto-fronts, outperforming BMOPSO. Besides the proper performance of DFGA, NSGA-II and SPEA2 in training phase, they do not converge to their optimal solutions in the same amount of time. DFGA converges faster than the rest, likely due to its deterministic initialization, which allows the algorithm to avoid evaluating solutions that are far from reaching the optimal value. Among the other two, NSGA-II converges faster than SPEA2, whose trail spreads across higher non-optimal $f_2(\boldsymbol{x})$ values. In contrast to the training performance of the aforementioned algorithms, BMOPSO does not show a suitable convergence. In fact, their final values are far from matching the Pareto-fronts of DFGA, NSGA-II or SPEA2. It is noteworthy that BMOPSO fronts keep high $f_2(\boldsymbol{x})$ values, which demonstrates that the algorithm has not been capable of improving solutions with a small number of channels. Owing to this behavior, it is not possible to assure that BMOPSO will reach proper performances in testing phase.

Considering these convergence results, it might be argued that meta-heuristics that work with discrete solutions (i.e., GA, BDE, DFGA, NSGA-II, SPEA2) show superior convergence results, which could be somewhat expected due to the nature of the problem. On the one hand, it can be said that local search strategies that rely on mutation, crossover and strength operators favor the conver-

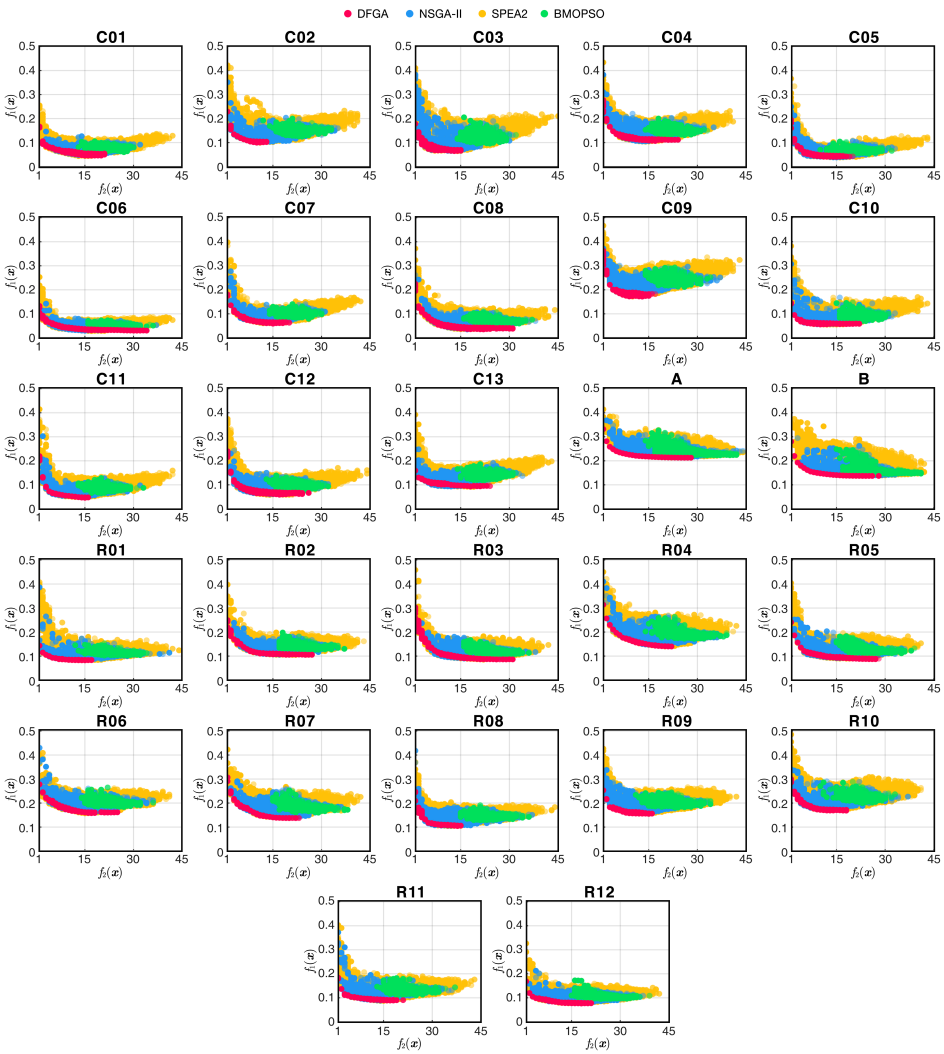


Figure 4: Evolution of Pareto-optimal solutions of the multi-objective meta-heuristics for each subject across all the repetitions: DFGA (red), NSGA-II (blue), SPEA2 (yellow) and BMOPSO (green).

gence in the P300-based BCI channel selection problem (Yang, 2014; Bozorg-Haddad et al., 2017; Zitzler et al., 2001). On the other hand, the behavior of BPSO and BMOPSO could imply that the discretization of continuous solutions cannot follow small value changes, hindering the local exploitation of the continuous-based algorithms if their hyperparameters have not been properly fixed.

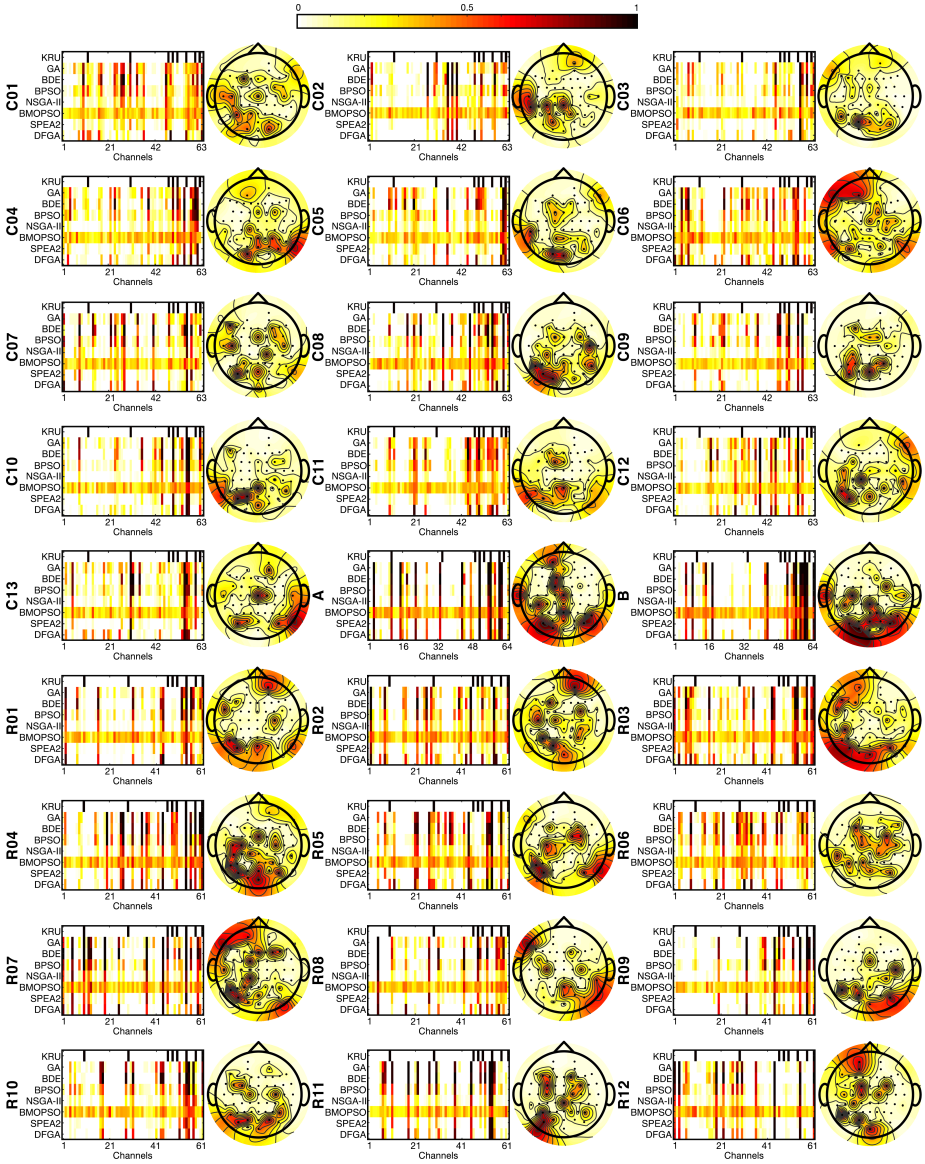


Figure 5: Channel ranks of the selected and the Pareto-optimal solutions for single-objective (GA, BDE, BPSO) and multi-objective (NSGA-II, BMOPSO, SPEA2, DFGA) meta-heuristics, respectively. Krusienski's 8-channel set (KRU) is also included. Averaged scalp distributions over the algorithms are depicted as well.

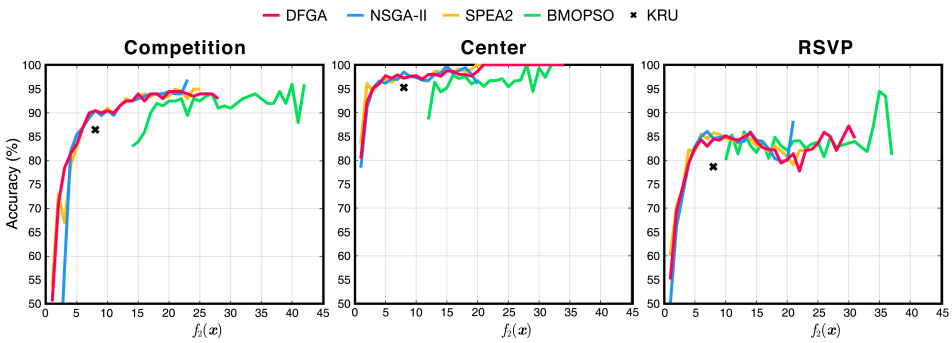


Figure 6: Testing character accuracies of the final Pareto-fronts returned by multi-objective meta-heuristics (DFGA, NSGA-II, SPEA2, BMOPSO) for the averaged subjects of each database. For comparison purposes, Krusienski's set (KRU) is also depicted.

5.2. Channel distributions

Averaged channel ranks of the figure 5 show that meta-heuristics in general have a slight tendency to mainly select electrodes over the occipital lobe. However, the optimal channel set is clearly different for each subject. This behavior confirms the fact that a customized channel selection procedure prior to the BCI session benefits the subsequent performance. Despite the Krusienski's (Krusienski et al., 2008) common 8-channel set is suitable as a general rule of thumb, results have not considered that combination optimal for any subject or database. This fact is reinforced in the testing phase, where both single-objective and multi-objective algorithm solutions have outperformed the 8-channel set, as can be noticed in the table 3 and the figure 6. Moreover, the computed Pareto-fronts have not generally kept solutions with more than $\sim$ 20 channels. This fact suggests that a small set of channels is able to reach similar or even better performances than the entire set, reducing the dimensionality and the computational cost of the BCI processing framework.

According to the previous analysis, meta-heuristics that converged faster for this optimization problem have also succeeded in finding the most relevant channels for each subject. As can be seen in the figure 5, GA, BDE, BPSO, DFGA, NSGA-II and SPEA2 reiteratively select a specific combination of channels, which is different for each subject. By contrast, BMOPSO do not show clear differences between channel ranks, which once again indicates a lack of convergence to a global optimum.

From the well-defined electrodes that were repeatedly selected for each subject, algorithms shown a special focus on the occipital cortex. From a biological point of view, this tendency is sound. As aforementioned, the $f_1(\boldsymbol{x})$ objective is aimed to maximize the classification performance between target and non-target event-related stimuli, elicited through a visual odd-ball task. The response is therefore modeled as an event-related potential (ERP) composed by several components, such as P1, N1, P2, N2 or P3; which are taken into account when extracting and classifying the features. Among them, P3 (i.e., P300) should be the most prominent one in the RCP (Wolpaw and Wolpaw, 2012; Polich, 2007). The primary visual cortex, highly specialized in processing information about visual stimuli, static and moving objects; is located at the posterior part of the occipital lobe (Standring, 2016). Hence, it is expected that occipital electrodes contain relevant discriminative information about target (i.e., ERP is present) and non-target signals (i.e., no ERP should be present) and thus, that they would likely be selected in the channel selection process. Nevertheless, the optimal channel sets are clearly different among subjects, which is relative common in the literature (Jin et al., 2010; Gonzalez et al., 2013; Kee et al., 2015; Chaurasiya et al., 2017). This fact should not be surprising, since classifiers are also optimized for each subject because of the inter-subject and inter-session variability of the EEG signals (Martínez-Cagigal et al., 2017). Even though the rationale behind the fact that optimal channel sets differ among subjects is not clear, it is believed that EEG is highly sensitive to external factors, such as inter-subject variations in the cap positions (Picton, 1992). In fact, it is common that EEG caps does not correctly fit some users, making some electrodes wobbly and producing noise. Furthermore, it should be taken into account that EEG channels cannot pinpoint neural sources owing to attenuation and volume conduction effects, being limited to a spatial resolution about 5 cm and hindering the location of the aforementioned sources in certain brain areas (Wolpaw and Wolpaw, 2012).

In short, we believe that most relevant channels for classification may not necessarily be the same among different users. As with feature selection and classification are calibrated for each subject, results show that the

channel selection procedure should be subject-optimized as well. Notwithstanding its usefulness as a preliminary approach, the common Kruskienski's set (Krusienski et al., 2008), which mainly locates channels over the parietal and occipital cortex, appears to be a suboptimal solution. It is noteworthy that our study is not intended to propose a general distribution of electrodes for any user, but to emphasize the need to customize the channel set for each subject.

5.3. Testing assessment

Single-objective approaches return a single solution with a specific number of channels, which minimizes the general objective $F(\boldsymbol{x})$. For this reason, averaged testing accuracies of table 3 should be taken into consideration together with the number of channels of each solution. According to the results, even though all the single-objective meta-heuristics have reached higher accuracies in comparison with the entire set and the Krusienski's 8-channel set, BDE stands out considering the channel-performance trade-off. BDE (competition: 92.0% with 14.5ch, center: 97.9% with 12.5ch, RSVP: 85.5% with 13.4ch) has reached the highest average accuracy with a scarce channel set, followed by GA (competition: 92.0% with 14.0ch, center: 97.4% with 12.4ch, RSVP: 84.6% with 13.4ch) and BPSO (competition: 92.0% with 14.0ch, center: 96.8% with 12.5ch, RSVP: 85.0% with 13.7ch). Nevertheless, all methods reach similar or even higher accuracies than the entire set of channels, probably due to the radical increase in dimensionality; outperforming as well the accuracy obtained by the typical 8-channel set. It is worthy to mention, however, that Krusienski's set also used less number of channels. In fact, the increase in testing accuracy of the three methods in comparison with the entire set (i.e., ALL) and the Krusienski's set (i.e., KRU) is statistically significant for almost all subjects (i.e., p–value< 0.05, Wilcoxon signed-rank test, false discovery rate corrected by the Benjamini-Hochberg procedure). In particular, the number of subjects (out of 27) that yielded significant differences were: 23 (ALL vs. GA), 27 (ALL vs. BDE), 22 (ALL vs. BPSO), 21 (KRU vs. GA, BDE or BPSO). As expected, the difference among GA, BDE and BPSO results is not significant. A detailed table with the p–values of each subject and comparison is included in the supplementary material. Therefore, it can be assured that GA, BDE and PSO outperform ALL and KRU; and that their solutions are similar in terms of reached accuracies.

The main advantage of the multi-objective meta-heuristics in comparison with the single-objective ones is that they return a set of optimal solutions for each number of channels, allowing the user to select the most appropriate configuration. In fact, not only these solutions indicate the number of channels that already reaches the maximum performance, but also their scarce solutions overcome the Krusienski's set. According to the figure 6, the typical 8-channel set is approximately outperformed using only 4 channels by DFGA, NSGA-II and SPEA2. These results are similar or even better than the individual solutions reached by single-objective algorithms. As can be noticed, DFGA, NSGA-II and SPEA2 reach similar performance results, which improve as the number of channels increase. Those results also outperform BMOPSO, whose solutions, in spite of using more channels, generally obtain lower accuracies. Results also show that there is a point for each subject where accuracies come to a standstill. In particular, using more than 15 channels in the competition or RSVP databases may be counter-productive; as well as using more than 20 channels in the center database. This fact reinforces the usefulness of dimensionality reduction techniques, such as channel and feature selection or classifier regularization, to assure a suitable testing performance in BCI systems.

Regarding previous studies, we consider that MI-based (Lv and Liu, 2008; Hasan and Gan, 2009; Hasan et al., 2010; Wei and Wang, 2011; Kee et al., 2015; Aler and Galván, 2015; Franklin Alex Joseph and Govindaraju, 2019; Zhang and Wei, 2019; González et al., 2019) and auditory potential (Gonzalez et al., 2014) BCI studies are not comparable in terms of performance, since those control signals are generally less reliable than P300 potentials and thus, obtain significantly lower accuracies. From the P300-based BCI studies, the reached accuracies of our work are similar or even higher than those reported previously. The most straightforward comparison comes from the III BCI Competition database (2 subjects), used by (Kee et al., 2015; Arican and Polat, 2020; Gonzalez et al., 2013; Perseh and Sharafat, 2012). Kee et al. (2015) reached an average accuracy of 93.6% with 22.3 channels using GA; and 94.9% with 25.7 channels using NSGA-II. Arican and Polat (2020) reached an averaged accuracy of 89.90% with 8 channels using BPSO and a boosted tree classifier. All of them used 15 sequences. A combination of wavelets and BPSO was also used by Perseh and Sharafat (2012), obtaining 85% with 31 channels; and Gonzalez et al. (2013), 67.5% with 33.5 channels using only 5 sequences. As can be seen, it is hard to compare the accuracies provided each study reported solutions with different number of channels or sequences. In our study, GA, BDE and BPSO yielded an averaged accuracy of 92% with 14 channels; and NSGA-II, SPEA2 and DFGA achieved 90% with only 7 channels, which increased until a maximum of 97% with 23 channels using 15 sequences. There are also studies with custom databases, such as Chaurasiya et al. (2017) (9 subjects, 15 sequences), who obtained a mean of 92.8% with 26.1 channels using MOBDE; or Jin et al. (2010) (11 subjects, 15 sequences), who tested a chinese-based RCP using PSO and LDA, reaching a mean of 71.09% with 7.63 channels. Besides the III BCI Competition database (2 subjects), our study also comprises the results with two additional databases: Center Speller (13 subjects, 10 sequences), and RSVP Speller (12 subjects, 10 sequences). However, no direct comparison can be made since there are no previous studies that have tested any meta-heuristic for selecting

channels with any paradigm apart from RCP. In terms of accuracy, our results for single-objective (center: 97.36% with 12.46 channels, RSVP: 85.03% with 13.5 channels) and multi-objective (center: 98.46% with 8 channels, RSVP: 85.73% with 8 channels) algorithms are similar to the performances reported in the literature (Treder et al., 2011; Acqualagna and Blankertz, 2013).

5.4. Hyperparameters

The main drawback of the majority of meta-heuristics rely on the need to fix hyperparameters, which usually depends on the context of the problem. Poorly chosen values may cause convergence issues, hindering the performance of the algorithm (Yang, 2014). In this context, the quality of a meta-heuristic must not be only assessed according to the performance results, but also taking into account the number of required hyperparameters. The less hyperparameters, the more probability to assure the reliability and generalization of the results. GA, DFGA, NSGA-II and SPEA2 only require mutation and crossover rates to be fixed. Fortunately, these parameters are widely studied in the literature, commonly fixed at $1/N$ for the mutation rate, and 0.90–0.95 for the crossover rate (Deb et al., 2002; Deb, 2005; Yang, 2014). A similar approach is followed in BDE, whose extra parameters are intended to perform a mutation procedure (Wang et al., 2012). BPSO and BMOPSO add three more hyperparameters (i.e., personal and global confidences and maximum velocity). Although there are several approaches that have tried to find global relations among their values, in light of the results, further endeavors should be made in order to make PSO algorithms problem-independent (Clerc and Kennedy, 2002; Yang, 2014). Since these hyperparameters directly weighs the velocity of the particles, which is then used as an input of a transfer function, care must be taken in order to limit their values in the range $[0, 1]$. Otherwise, particles will not tend to improve their solutions, limiting the global and local exploitation. In this study, we have fixed the hyperparameters according to the literature, as indicated in the table 1. These values have yield suitable performances, but could have been improved by means of a hyperparameter optimization or following an adaptive approach. It is also worthy to mention that the weights of the single-objective approach to merge both objectives were heuristically set to ω_1 and ω_2 to 0.7 and 0.3, respectively, in view or preliminary results (Martínez-Cagigal and Hornero, 2017a,b; Martínez-Cagigal et al., 2018). Note that the difference of values of ω_1 and ω_2 would cause a strengthen of solutions in a certain $f_2(\boldsymbol{x})$ range, while avoiding the search in other spaces. The supervisor could vary the ω_1/ω_2 ratio to obtain different optimal solutions, simulating the search over the $f_2(\boldsymbol{x})$ spectrum as multi-objective approaches do.

5.5. Computational cost

A comparison between computational costs of different meta-heuristics is tricky, forcing to consider several aspects at the same time. On the one hand, the table 2 details the approximate duration of a generation and the number of evaluations that comprised a generation. Note that the number of evaluations per generation differs depending on the meta-heuristic strategy and thus, algorithms can only be compared in terms of the duration of a single evaluation. For this reason, the number of generations of each algorithm has been adapted in order to assure a fair comparison among them, so that every single method performs a total of 4000 evaluations. On the other hand, it is also essential to consider further issues, such as the convergence of the algorithms, their search depth and the programming approach. When abstracting a meta-heuristic as a black box, the total time of the execution varies according to the required number of generations to reach a suitable convergence. These differences usually affect the computational cost in a higher extent than the individual duration of a evaluation, making it essential to be taken into consideration. Moreover, a correct implementation of these meta-heuristics should employ a hash table to match the previously computed solutions with their fitnesses. Note that an intense search depth will inevitable generate repeated solutions. The hash map acts as a remainder and allows avoiding unnecessary evaluations. Hence, it is noteworthy that the computation time of the generations declines exponentially when the algorithm goes on. Note that table 2 measurements were made without considering any hash table (i.e., an initial generation).

The overall complexity of DFGA in asymptotic notation behaves as $O(N_o N^2)$, where N_o is the number of objectives. An analysis of the complexity of each DFGA step is detailed in the supplementary material, which demonstrates that the exponential increase in operations is mainly due to the dual-front sorting procedure. However, this trend is similar to other recent multi-objective algorithms, such as NSGA-II, SPEA2 and BMOPSO, whose complexity behaves as $O(N_o m^2)$ (Tripathi et al., 2007; Curry and Dagli, 2014a). Note that m indicates the population size which, in case of DFGA, equals to N (the no. channels) due to the deterministic initialization. This fact partly explains why DFGA performs a higher number of evaluations in a single generation. It is also noteworthy that the presented asymptotic complexity analysis involves only the advancing of a single generation, allowing the comparison among different algorithms, since their convergences are not deterministic (Curry and Dagli, 2014a). Moreover, these complexities indicates the worst cases, which usually decrease as generations increase due to the hash table implementation.

According to the table 2, NSGA-II and DFGA are the least time-consuming, spending less than 600 ms per evaluation with the selected hyperparameters. In addition, they have demonstrated excellent convergence abil-

ities, making them excellent multi-objective approaches to this problem. By contrast, BPSO and BMOPSO not only are the most time-consuming algorithms, but also their performances are inferior. For the single-objective approaches, it is worthwhile to use GA or BDE, whose convergence abilities balance out their evaluation costs. In any case, the overall duration of these algorithms restrict their application to the calibration session, where the weights of the classifier are optimized for each subject. Then, the selected channels should be further applied in the testing sessions.

5.6. Guidelines

A series of guidelines or practical recommendations for the application of meta-heuristics to BCI systems may be derived from the discussed results:

1. Multi-objective algorithms should be used instead of single-objective meta-heuristics if computation time is not an issue. Otherwise, it is preferable even using deterministic algorithms, such as BE, to provide sub-optimal but acceptable solutions.

2. Discrete algorithms that use mutation, crossover or strength operators should be preferred (e.g., single-objective: GA, BDE; multi-objective: DFGA, NSGA-II, SPEA2).

3. Whether discretization is required to adapt a continuous-based meta-heuristic to the BCI framework, avoid using transfer functions and attempt to redefine the equations (section 3.5). Still, if conversion via transfer functions is used, care should be taken with hyperparameters values. Assure that the input of the function always lies within the range $[0, 1]$. Otherwise, the probability of change of the solutions would drastically increase, hindering the convergence of the algorithm. Furthermore, distance metrics should not be employed after applying the discrete transformation unless it is the Hamming distance.

4. For single-objective algorithms, use an aggregation approach to minimize two objectives at one: number of channels and performance error (section 3.2). AUC-based modeling of the performance should be preferred instead of accuracies in order to increment the resolution of the objective values.

5. Multi-objective repository limitation strategies, such as crowding distances or distance sought, are not necessary in the BCI channel selection problem and should be avoided to prevent worthless computational costs (section 3.5).

6. A hybrid meta-heuristic that also employs deterministic methods, such as DFGA, is preferred. DFGA has reached similar accuracies than NSGA-II and SPEA2, but converges faster (section 4).

7. A convergence detection method to stop the iterations is recommended rather than using a maximum generation limit for practical purposes (e.g., none or petty changes along the n last generations).

8. Repeated solutions across generations are unavoidable. Make sure to implement a hash map (e.g., dictionary) for matching previously computed solutions with their fitnesses, in order to avoid further unnecessary evaluations.

5.7. Contributions, limitations and future endeavors

According to the experimental outcomes, it has been demonstrated the utility of meta-heuristics to find an optimal combination of channels in P300-based BCI systems. The importance of selecting an optimal channel set for each user has been highlighted as well. Moreover, to the best of our knowledge, this is the first study that provides a comprehensive comparison of different meta-heuristics that can be applied to the BCI channel selection problem. Previous studies have isolated the application of BPSO, BDE, GA and NSGA-II, but no comparison has been performed; whereas this manuscript has included a total of 3 single-objective (i.e., GA, BDE, BPSO) and 4 multi-objective (i.e., DFGA, NSGA-II, SPEA2, BMOPSO) algorithms. As a result, GA, BDE, BPSO, DFGA, NSGA-II and SPEA2 have been reached high performances in testing phase, outperforming the entire set and the common Krusienski's 8-channel set in three databases with different paradigms. Due to the characteristics of the BCI framework, none of the well-known methods can be applied in a productive way without a proper modification. For that reason, DFGA, a new multi-objective meta-heuristic, has been especially designed to optimize channel or feature sets in BCI systems. In this way, the supervisor may apply DFGA in the first session, select an appropiate channel set and avoid placing extra electrodes for the next BCI sessions. Furthermore, in order to ease the application of meta-heuristics in these systems, a series of guidelines have been detailed. To sum up, the main contributions of this study are the following: (i) proposal of a novel multi-objective algorithm, DFGA, especially designed for this problem, (ii) comparison of a total of 7 meta-heuristics to the P300-based BCI channel selection problem, (iii) enumeration of a detailed set of guidelines to adapt any meta-heuristic for the channel selection, and (iv) successful evaluation with three databases that employ different P300-based paradigms.

Despite the aforementioned strengths, several limitations can be pointed out. Firstly, since the purpose of the manuscript was focused on the channel selection procedure, only basic feature extraction (i.e., down-sampling) and classification (i.e., LDA) methods have been applied. Testing accuracies, especially those than belong to crowded channel sets, could have been improved by using regularization techniques (Perseh and Sharafat, 2012; Tahernezhad-Javazm et al., 2017; Blankertz et al.,

2011). Moreover, the algorithms entail high computational costs. Further endeavors should be aimed at assigning stopping criteria that could avoid the computation of worthless generations, allowing a better estimation of the total duration for each model. The computational cost is mainly caused by the wrapper nature of the algorithms, which evaluate the quality of a solution by training and testing different LDA models (Saeys et al., 2007). Embedded techniques (e.g., heuristic search methods), which look for optimal sets inside the classifier constructor, are less intensive than wrappers (Guyon and Elisseeff, 2003; Saeys et al., 2007). A future endeavor could be aimed at developing new embedded techniques that could reduce the computational cost by modifying the training procedures of certain classifiers. In addition, as it have been seen for some outlier subjects, the use of this meta-heuristics could not improve the system's performance which, together with the computational burden, could be viewed as a waste of time. Care should be also taken when using transfer functions, such as those used in BPSO or BMOPSO, since they could be fruitless to the proper exploitation of the discrete space. Hyperparameters were fixed according to the recomendations of the literature. However, an optimization of these values would be beneficial to the final performance of the algorithms. Adaptive approaches that vary the hyperparameters in function of the generation could also enhance the results. It should be also mentioned that the competition database contains more training trials than those that are commonly in practice. Another future research line could be focused on assessing the performance of these methods with less training trials. Finally, although it has not been explored in study, results suggest that the proposed meta-heuristics could be also applied to feature selection problems.

6. Conclusions

A comprehensive comparison among 7 different meta-heuristics applied to the P300-based BCI channel selection problem has been performed in this study. In particular, 3 single-objective and 4 multi-objective algorithms have been included. Due to the discrete characteristics of the BCI framework, the majority of them have been modified in different ways in order to adapt them to the aforementioned problem. For this reason, a series of guidelines or practical recommendations have been detailed as an aid for further adaptations. A novel multi-objective algorithm has been especially developed for BCI systems: DFGA. Methods have been tested with three public databases that used different stimulation paradigms: competition (2 users with 64ch., RCP), center speller (13 users with 63ch., CS paradigm) and RSVP speller (12 users with 61ch., RSVP). Results show that meta-heuristics are able to provide solutions that simultaneously use few number of channels and reach high accuracies. In fact, the entire set of channels and the common Krusienki's 8-channel set have been outperformed by the 7 methods, demonstrating their usefulness to provide an optimized channel set for each user.

The main findings of the study can be summarized as follows:

1. Optimal channel sets show a high inter-subject variability, which makes essential the optimization for each individual, instead of using a common set.

2. Inherently discrete algorithms (i.e., GA, BDE, DFGA, NSGA-II, SPEA2) reach higher performances due to the dichotomous nature of the problem.

3. Among single-objective meta-heuristics, GA, and BDE provide suitable convergences and high accuracies. Regarding multi-objective algorithms, DFGA, NSGA-II and SPEA2 provided competitive results.

4. A balanced combination of deterministic and stochastic techniques is beneficial. DFGA reaches an excellent performance, as well as NSGA-II and SPEA2, but converges considerably faster to their optimal solutions.

5. Hyperparameter tuning is crucial. BMOPSO could not converge to an optimal solution, whereas it is possible to guarantee the convergence of the rest in a single run.

Acknowledgments

This study was partially funded by projects DPI2017-84280-R of 'Ministerio de Ciencia, Innovación y Universidades' and 'European Regional Development Fund' (FEDER), 'Análisis y correlación entre el genoma completo y la actividad cerebral para la ayuda en el diagnóstico de la enfermedad de Alzheimer' (Cooperation Programme Interreg V-A Spain-Portugal POCTEP 2014-2020) of the European Commission and FEDER, and by 'CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN)' through 'Instituto de Salud Carlos III' cofunded with FEDER. V. Martínez-Cagigal was in receipt of a PIF-UVa grant of the University of Valladolid. The authors declare no conflicts of interest.

References

Acqualagna, L., Blankertz, B., 2013. Gaze-independent BCI-spelling using rapid serial visual presentation (RSVP). Clinical Neurophysiology 124 (5), 901–908.
URL http://dx.doi.org/10.1016/j.clinph.2012.12.050

Aler, R., Galván, I. M., 2015. Optimizing the number of electrodes and spatial filters for Brain-Computer Interfaces by means of an evolutionary multi-objective approach. Expert Systems with Applications 42 (15-16), 6215–6223.
URL http://dx.doi.org/10.1016/j.eswa.2015.03.008

Arican, M., Polat, K., 2020. Binary particle swarm optimization (BPSO) based channel selection in the EEG signals and its application to speller systems. Journal of Artificial Intelligence and Systems 2 (1), 27–37.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal statistical society: series B (Methodological) 57 (1), 289–300.

Blankertz, B., Lemm, S., Treder, M., Haufe, S., Müller, K. R., 2011. Single-trial analysis and classification of ERP components - A tutorial. NeuroImage 56 (2), 814–825.
URL http://dx.doi.org/10.1016/j.neuroimage.2010.06.048

Blankertz, B., Müller, K.-R., Krusienski, D. J., Schalk, G., Wolpaw, J. R., Schlögl, A., Pfurtscheller, G., Millán, J. D. R., Schröder, M., Birbaumer, N., 2006. The BCI competition III: Validating alternative approaches to actual BCI problems. IEEE Transactions on Neural Systems and Rehabilitation Engineering 14 (2), 153–159.

Bozorg-Haddad, O., Solgi, M., A. Loaiciga, H., 2017. Meta-Heuristic and Evolutionary Algorithms for Engineering Optimization.
URL www.wiley.com

Cecotti, H., Rivet, B., Congedo, M., Jutten, C., Bertrand, O., Maby, E., Mattout, J., 2011. A robust sensor-selection method for P300 brain-computer interfaces. Journal of Neural Engineering 8 (1), 16001.
URL http://www.ncbi.nlm.nih.gov/pubmed/21245524

Chaurasiya, R. K., Londhe, N. D., Ghosh, S., 2017. Multi-objective binary DE algorithm for optimizing the performance of Devanagari script-based P300 speller. Biocybernetics and Biomedical Engineering 37 (3), 422–431.
URL http://dx.doi.org/10.1016/j.bbe.2017.04.006

Clerc, M., Kennedy, J., 2002. The Particle Swarm–Explosion, Stability, and Convergence in a Multidimensional Complex Space. IEEE Transactions on Evolutionary Computation 6 (1), 58–73.

Coello, C. a., Reyes-Sierra, M., 2006. Multi-Objective Particle Swarm Optimizers: A Survey of the State-of-the-Art. International Journal of Computational Intelligence Research 2 (3), 287–308.
URL http://www.ijcir.com/publishedPapers.php?showDetails=Y{&}idArticle=68{&}volume=1{&}number=1{&}volume{_}id=2{&}number{_}id=4

Colwell, K. A., Ryan, D. B., Throckmorton, C. S., Sellers, E. W., Collins, L. M., 2014. Channel selection methods for the P300 Speller. Journal of Neuroscience Methods 232, 6–15.

Curry, D. M., Dagli, C. H., 2014a. Computational complexity measures for many-objective optimization problems. Procedia Computer Science 36 (C), 185–191.
URL http://dx.doi.org/10.1016/j.procs.2014.09.077

Curry, D. M., Dagli, C. H., 2014b. Computational complexity measures for many-objective optimization problems. Procedia Computer Science 36, 185–191.

Deb, K., 2005. Multi-Objective Optimization.
URL http://link.springer.com/10.1007/0-387-28356-0{_}10

Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation 6 (2), 182–197.

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

Franklin Alex Joseph, A., Govindaraju, C., 2019. Channel selection using glow swarm optimization and its application in line of sight secure communication. Cluster Computing 22 (s5), 10801–10808.

Gonzalez, A., Nambu, I., Hokari, H., Iwahashi, M., Wada, Y., 2013. Towards the classification of single-trial event-related potentials using adapted wavelets and particle swarm optimization. Proceedings - 2013 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2013, 3089–3094.

Gonzalez, A., Nambu, I., Hokari, H., Wada, Y., 2014. EEG Channel Selection Using Particle Swarm Optimization for the Classification of Auditory Event-Related Potentials. Scientific World Journal 2014, 350270.

González, J., Ortega, J., Damas, M., Martín-Smith, P., Gan, J. Q., 2019. A new multi-objective wrapper method for feature selection – Accuracy and stability analysis for BCI. Neurocomputing 333, 407–418.

Guyon, I., Elisseeff, A., 2003. An Introduction to Variable and Feature Selection. Journal of Machine Learning Research (JMLR) 3 (3), 1157–1182.

Hasan, B. A. S., Gan, J. Q., 2009. Multi-Objective Particle Swarm Optimization for Channel Selection in Brain-Computer Interfaces. The UK Workshop on Computational Intelligence, 2–7.
URL http://www.abdn.ac.uk/{%}7B{~}{%}7Dmbi361/papers/bashar{_}UKCI09.pdf

Hasan, B. A. S., Gan, J. Q., Zhang, Q., 2010. Multi-objective evolutionary methods for channel selection in brain-computer interfaces: some preliminary experimental results. IEEE Congress on Evolutionary Computation (CEC-2010), 1–6.

Holland, J. H., 1992. Genetic algorithms. Scientific american 267 (1), 66–73.

Jin, J., Allison, B. Z., Brunner, C., Wang, B., Wang, X., Zhang, J., Neuper, C., Pfurtscheller, G., 2010. P300 Chinese input system based on Bayesian LDA. Biomedizinische Technik 55 (1), 5–18.

Jobson, J. D., 1991. Applied multivariate data analysis: volume I: Regression and Experimental Design. Springer.

Kee, C.-Y., Ponnambalam, S. G., Loo, C.-K., 2015. Multi-objective genetic algorithm as channel selection method for P300 and motor imagery data set. Neurocomputing 161, 120–131.
URL http://linkinghub.elsevier.com/retrieve/pii/S0925231215002295{%}5Cnhttp://www.scopus.com/inward/record.url?eid=2-s2.0-84929047987{&}partnerID=tZOtx3y1

Kennedy, J., Eberhart, R., 1995. Particle swarm optimization. Neural Networks, 1995. Proceedings., IEEE International Conference on 4, 1942—-1948 vol.4.

Kennedy, J., Eberhart, R. C., 1997. A Discrete Binary Version of the Particle Swarm Algorithm. 1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation 5, 4–8.

Krusienski, D. J., Sellers, E. W., McFarland, D. J., Vaughan, T. M., Wolpaw, J. R., 2008. Toward enhanced P300 speller performance. Journal of Neuroscience Methods 167 (1), 15–21.

Lipowski, A., Lipowska, D., 2012. Roulette-wheel selection via stochastic acceptance. Physica A: Statistical Mechanics and its Applications 391 (6), 2193–2196.

Lv, J., Liu, M., 2008. Common Spatial Pattern and Particle Swarm Optimization for Channel Selection in BCI. 2008 3rd International Conference on Innovative Computing Information and Control, 457.
URL http://ieeexplore.ieee.org/xpls/abs{_}all.jsp?arnumber=4603646{%}5Cnhttp://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4603646

Martínez-Cagigal, V., Gomez-Pilar, J., Álvarez, D., Hornero, R., 2017. An asynchronous P300-based Brain–Computer Interface web browser for severely disabled people. IEEE Transactions on Neural Systems and Rehabilitation Engineering 25 (8), 1332–1342.

Martínez-Cagigal, V., Hornero, R., 2017a. A Binary Bees Algorithm for P300-Based Brain-Computer Interfaces Channel Selection. In: Advances in Computational Intelligence. IWANN 2017. Lecture Notes in Computer Science, 1st Edition. Springer International Publishing AG, Cádiz, Spain, pp. 453–463.
URL https://link.springer.com/chapter/10.1007/978-3-319-59147-6{_}39

Martínez-Cagigal, V., Hornero, R., 2017b. P300-based Brain-Computer Interface channel selection using swarm intelligence. Revista Iberoamericana de Automática e Informática Industrial 14 (4), 372–383.

Martínez-Cagigal, V., Santamaría-Vázquez, E., Hornero, R., 2018. A novel hybrid swarm algorithm for P300-based BCI channel selection. In: IFMBE Proceedings. Vol. 68.

Mirjalili, S., Lewis, A., 2013. S-shaped versus V-shaped transfer functions for binary Particle Swarm Optimization. Swarm and Evolutionary Computation 9, 1–14.
URL http://dx.doi.org/10.1016/j.swevo.2012.09.002

Perseh, B., Sharafat, A. R., jun 2012. An Efficient P300-based BCI Using Wavelet Features and IBPSO-based Channel Selection. Journal of

Medical Signals and Sensors 2 (3), 128–143.
Picton, T. W., 1992. The P300 wave of the human event-related potential. Journal of Clinical Neurophysiology 9 (4), 456–479.
Poli, R., Kennedy, J., Blackwell, T., 2007. Particle swarm optimization. Swarm Intelligence 1 (1), 33–57.
URL http://link.springer.com/10.1007/s11721-007-0002-0
Polich, J., 2007. Updating P300: An Integrative Theory of P3a and P3b. Clin. Neurophysiol. 118 (10), 2128–2148.
Reyes-Sierra, M., Coello, C. A., 2005. Improving PSO-Based Multi-objective Optimization Using Crowding, Mutation and E-Dominance. Lecture Notes in Computer Science 3410, 505–519.
URL http://link.springer.com/10.1007/978-3-540-31880-4{_}35
Saeys, Y., Inza, I., Larrañaga, P., 2007. A review of feature selection techniques in bioinformatics. Bioinformatics 23 (19), 2507–2517.
Standring, S., 2016. Gray's anatomy: the anatomical basis of clinical practice, 41st Edition. Elsevier Ltd, London.
Storn, R., Price, K., 1997. Differential Evolution - A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. Journal of Global Optimization 11 (4), 341–359.
Tahernezhad-Javazm, F., Azimirad, V., Shoaran, M., 2017. A review and experimental study on application of classifiers and evolutionary algorithms in EEG based brain-machine interface systems. Journal of Neural Engineering.
URL http://iopscience.iop.org/article/10.1088/1741-2552/aa8063
Treder, M. S., Schmidt, N. M., Blankertz, B., 2011. Gaze-independent brain-computer interfaces based on covert attention and feature attention. J Neural Eng 8 (6), 66003.
URL http://dx.doi.org/10.1088/1741-2560/8/6/066003
Tripathi, P. K., Bandyopadhyay, S., Pal, S. K., 2007. Multi-Objective Particle Swarm Optimization with time variant inertia and acceleration coefficients. Information Sciences 177 (22), 5033–5049.
Wang, L., Fu, X., Mao, Y., Ilyas Menhas, M., Fei, M., 2012. A novel modified binary differential evolution algorithm and its applications. Neurocomputing 98, 55–75.
URL http://dx.doi.org/10.1016/j.neucom.2011.11.033
Wei, Q., Wang, Y., 2011. Binary Multi-Objective Particle Swarm Optimization for Channel Selection in Motor Imagery Based Brain-Computer Interfaces. 2011 4th International Conference on Biomedical Engineering and Informatics (BME I), 667–670.
Wolpaw, J., Wolpaw, E. W., 2012. Brain-computer interfaces: principles and practice. OUP USA.
Yang, X.-S., 2014. Nature-Inspired Optimization Algorithms, 1st Edition. Elsevier Inc.
Yu, L., Liu, H., 2003. Feature Selection for High-Dimensional Data: A Fast Correlation-Based Filter Solution. Proceedings, Twentieth International Conference on Machine Learning 2, 856–863.
Zhang, L., Wei, Q., 2019. Channel selection in motor imaginary-based brain-computer interfaces: a particle swarm optimization algorithm. Journal of integrative neuroscience 18 (2), 141–152.
Zitzler, E., Laumanns, M., Thiele, L., 2001. SPEA2: Improving the Strength Pareto Evolutionary Algorithm. Tech. rep., ETH Zurich, Zurich.
Zou, H., Hastie, T., 2005. Erratum: Regularization and variable selection via the elastic net (Journal of the Royal Statistical Society. Series B: Statistical Methodology (2005) 67 (301-320)). Journal of the Royal Statistical Society. Series B: Statistical Methodology 67 (5), 768.

Supplementary material

For the sake of replicability and completeness, additional methodological details and further analyses are included in this supplementary material.

6.1. Computational complexity of DFGA

In this section, the computational complexity of the proposed multi-objective algorithm, DFGA, is analyzed in terms of the asymptotic notation (i.e., *Big* O notation). In order to present a detailed estimate of the complexity, each step of the pseudo-code (see figure 2) is analyzed in the following lines:

0. **Backward elimination (BE).** This step is executed once, since it encompasses the deterministic initialization of DFGA. Therefore, its complexity entirely depends on the size of the solution (i.e., chromosome) – that is, the number of channels, N. BE sequentially removes one channel at a time. Hence, it is straightforward that the complexity of the first iteration would be $O(1)$ (checking the $f_1(\boldsymbol{x})$ of the entire set of channels). Then, the complexity of the second iteration would be $O(N)$, which linearly decreases to $O(N-(N-2)) = O(2)$ for the N-th iteration. The aggregated complexity would be $O(N^2 - N - \sum_{i=2}^{N-2} i)$.

1. **Dual-front sorting and parent selection.** The dual-front sorting divides the current repository into the optimal $\mathcal{O}$ and sub-optimal $\mathcal{S}$ sets. In order to do that, the algorithm first needs to check the dominance of every single combination of couples of solutions. This equals to the double of the binomial coefficient, since dominance must be checked in both directions: $O(2N_o \cdot \binom{N}{2}) = O(N_o(N^2 - N))$, where N_o is the number of objectives (in this case, $N_o = 2$). Note that, in DFGA, the size of the repository is equal to the number of channels N. Afterward, binary tournament selection is applied to select the parent population from the repository, adding a complexity of $O(N)$. Hence, the aggregated complexity of this step would be $O(N_o N^2 + N(1 - N_o))$.

2. **Genetic operators.** Since the length of the parent population is always N, single-point crossover requires $O(N)$ operations. By contrast, bit-string mutation requires $O(N^2)$ operations. Therefore, the aggregated complexity would be $O(N^2 + N)$.

3. **Synthetic solutions.** The generation of synthetic solutions is used to maintain a similar exploitation over the entire repository spectrum. The computation involves $N - 1$ roulette wheel selections. This well-known algorithm usually takes $O(\log N)$ operations, leading to an aggregated complexity of $O((N - 1)\log N)$. However, the complexity of the roulette wheel selection can be reduced to $O(1)$ by means of an "alias method" implementation, leading to an aggregate complexity of $O(N - 1)$ (Lipowski and Lipowska, 2012).

4. **Repository updating.** In order to update the repository, fitnesses of the last repository, the parent population and the synthetic solutions must be checked, leading to a complexity of $O(N + N + N - 1) = O(3N - 1)$.

The computational complexity of DFGA would be $O((2+N_o)N^2+(5-N_o)N-2-\sum_{i=2}^{N-2} i)$ including the initialization, or $O((1+N_o)N^2+(6-N_o)N-2)$ if a generic generation is considered. Therefore, the overall complexity of the algorithm behaves as $O(N_oN^2)$, influenced mainly by the dual-front sorting procedure. However, care must be taken when interpreting this analysis, since the complexity presented here involves the advancing of a single generation, not the algorithm's aggregated complexity. In this way, evolutionary algorithms can be compared when their convergence is not deterministic (Curry and Dagli, 2014b). It is noteworthy that $O(N_oN^2)$ complexity denotes the worst case, which usually decreases as the generations increase if a hash-map implementation is used.

6.2. Statistical analysis

The table 4 details the averaged accuracy results and *p*–values for the single-objective meta-heuristics results. Note that accuracies were averaged across the 20 repetitions of the algorithms for each subject and that *p*–values were obtained using a Wilcoxon signed-rank test. Then, the False Discovery Rate (FDR) was corrected using the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995).

6.3. Individual convergence curves

The figure 7 shows the individual convergence curves for the single-objective meta-heuristics for each subject, which where omitted in the original paper for saving space.

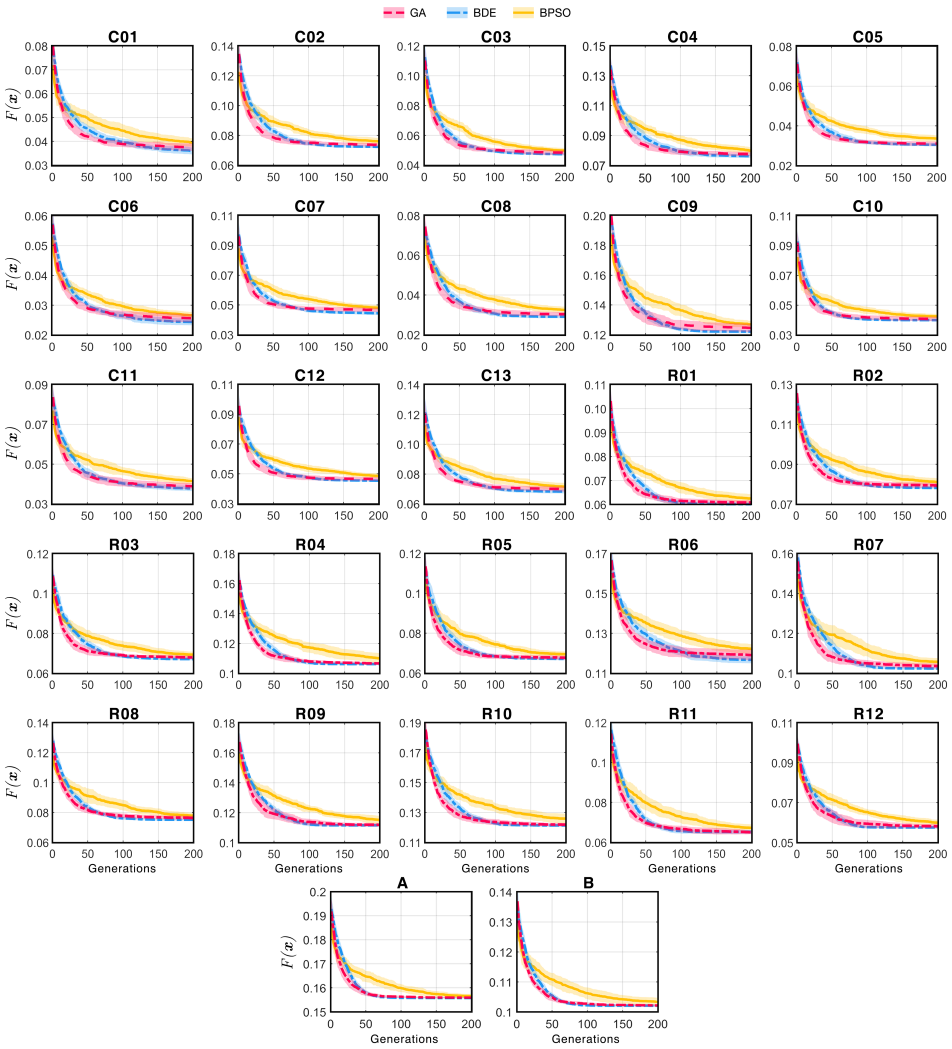


Figure 7: Individual convergence curves of single-objective meta-heuristics (GA, BDE and BPO) for each database in function of the $F(\boldsymbol{x})$ aggregated function. Mean values are displayed with solid lines, whereas the 95% confidence interval of the subjects' repetitions is indicated by the shaded area.

Table 4: Averaged accuracy results and p–values for the single-objective meta-heuristics.

Subject		GA		BDE		PSO		ALL	p–values of ALL vs.			KRU	p–values of KRU vs.			p–values of		
		Acc. (%)	N	Acc. (%)	N	Acc. (%)	N	Acc. (%)	GA	BDE	BPSO	Acc. (%)	GA	BDE	BPSO	GA vs. BDE	GA vs. BPSO	BDE vs. BPSO
Comp	A	90.35	14.4	90.4	14.2	89.55	14.15	94	**$8.2{\cdot}10^{-5}$**	**$8.2{\cdot}10^{-5}$**	**$8.2{\cdot}10^{-5}$**	84	**$8.2{\cdot}10^{-5}$**	**$8.2{\cdot}10^{-5}$**	**$8.2{\cdot}10^{-5}$**	0.9469	0.10205	0.045955
	B	94.1	14.2	93.8	14.35	94.2	14.75	90	**$8.2{\cdot}10^{-5}$**	**$8.2{\cdot}10^{-5}$**	**$8.2{\cdot}10^{-5}$**	89	**$8.2{\cdot}10^{-5}$**	**$8.2{\cdot}10^{-5}$**	**$8.2{\cdot}10^{-5}$**	0.5708	0.75442	0.44604
Center	U01	95.12	14.05	95.71	14	95.71	14.55	85.71	**$8.3{\cdot}10^{-5}$**	**$8.3{\cdot}10^{-5}$**	**$8.3{\cdot}10^{-5}$**	88.1	**0.000193**	**0.000193**	**0.000193**	0.86328	1	1
	U02	92.11	11.95	93.95	10.3	90.26	12.6	76.32	**$8.4{\cdot}10^{-5}$**	**$8.3{\cdot}10^{-5}$**	**$8.4{\cdot}10^{-5}$**	86.84	**0.000384**	**0.000193**	**0.003286**	0.5698	0.62207	0.021434
	U03	100	10.15	100	9.4	100	10.35	94.59	**$2.5{\cdot}10^{-5}$**	**$2.5{\cdot}10^{-5}$**	**$2.5{\cdot}10^{-5}$**	97.3	**$5{\cdot}10^{-5}$**	**$5{\cdot}10^{-5}$**	**$5{\cdot}10^{-5}$**	1	1	1
	U04	99.77	13.65	99.77	13.3	99.65	14.1	93.02	**$4.2{\cdot}10^{-5}$**	**$4.2{\cdot}10^{-5}$**	**$5.3{\cdot}10^{-5}$**	100	0.60938	0.60938	0.33621	1	1	1
	U05	97.93	12.05	97.44	12.05	98.05	11.7	85.37	**$8.3{\cdot}10^{-5}$**	**$8.3{\cdot}10^{-5}$**	**$8.3{\cdot}10^{-5}$**	90.24	**0.000193**	**0.000193**	**0.000193**	0.86328	1	0.58398
	U06	100	13.35	100	13.3	100	13.75	96.88	**$2.5{\cdot}10^{-5}$**	**$2.5{\cdot}10^{-5}$**	**$2.5{\cdot}10^{-5}$**	100	1	1	1	1	1	1
	U07	100	13.55	100	13.2	99.88	13	88.37	**$2.5{\cdot}10^{-5}$**	**$2.5{\cdot}10^{-5}$**	**$3.6{\cdot}10^{-5}$**	95.35	**$5{\cdot}10^{-5}$**	**$5{\cdot}10^{-5}$**	**$6.6{\cdot}10^{-5}$**	1	1	1
	U08	100	14.6	100	13.6	100	14.65	94.87	**$2.5{\cdot}10^{-5}$**	**$2.5{\cdot}10^{-5}$**	**$2.5{\cdot}10^{-5}$**	100	1	1	1	1	1	1
	U09	85.22	11.35	87.39	10.3	84.78	12.2	63.04	**$8.3{\cdot}10^{-5}$**	**$8.3{\cdot}10^{-5}$**	**$8.3{\cdot}10^{-5}$**	86.96	**0.004328**	0.55114	**0.001253**	0.046021	1	0.021434
	U10	95.31	12	94.8	11.25	96.02	11.55	91.84	**0.000103**	**$9{\cdot}10^{-5}$**	**$8.3{\cdot}10^{-5}$**	95.92	0.29687	**0.009141**	1	0.86328	0.89375	0.023275
	U11	97.88	12.95	98.13	12.45	99.13	12.95	77.5	**$8.3{\cdot}10^{-5}$**	**$8.3{\cdot}10^{-5}$**	**$8.3{\cdot}10^{-5}$**	100	**0.001058**	**0.001904**	**0.046875**	1	0.41608	0.39482
	U12	100	13.25	99.73	12	99.19	13.4	94.59	**$2.5{\cdot}10^{-5}$**	**$4.2{\cdot}10^{-5}$**	**$8.3{\cdot}10^{-5}$**	97.3	**$5{\cdot}10^{-5}$**	**0.000108**	**0.000298**	0.92857	0.40625	0.39482
	U13	99.23	13.45	98.72	11.85	98.59	13.25	82.05	**$7.8{\cdot}10^{-5}$**	**$8.3{\cdot}10^{-5}$**	**$8.3{\cdot}10^{-5}$**	100	0.090278	**0.006348**	**0.006348**	0.86328	0.89375	1
RSVP	R01	86.46	11.75	85.85	11.25	87.68	11.35	82.93	**0.000325**	**0.00061**	**0.00031**	85.37	**0.026953**	**0.44397**	**0.000863**	1	0.34553	0.11133
	R02	82.7	14.15	83.65	13.5	82.7	14.3	75.68	**0.000411**	**0.000258**	**0.000258**	78.38	**0.000507**	**0.000143**	**0.000505**	1	1	0.72518
	R03	82.69	15.6	84.74	14.55	82.95	15.8	82.05	**0.4458**	**0.003871**	0.31901	84.62	**0.017881**	0.89197	0.08385	0.15967	1	0.28538
	R04	72.83	16.4	73.04	16.35	70.76	16.85	71.74	0.22258	**0.048516**	0.20669	54.35	**0.000143**	**0.000143**	**0.000143**	1	0.34553	0.11133
	R05	92.64	13.75	92.36	14	92.5	14.35	91.67	0.19429	0.31901	0.31901	91.67	0.17	0.3284	0.3284	1	1	0.85212
	R06	84.21	15.5	83.82	15.35	84.08	14.8	76.32	**0.000258**	**0.000258**	**0.000411**	76.32	**0.000143**	**0.000143**	**0.000257**	1	1	0.85212
	R07	71	15.25	70.9	15.55	73.4	16.25	68	**0.005708**	**0.000449**	**0.000833**	60	**0.000143**	**0.000143**	**0.000143**	1	0.34553	0.11133
	R08	93.42	11.4	92.76	10.65	93.82	11.5	94.74	0.050875	**0.000258**	0.15885	84.21	**0.000143**	**0.000143**	**0.000143**	1	1	0.28538
	R09	88.1	11.25	88.1	11.1	87.86	12	73.81	**0.000258**	**0.000258**	**0.000258**	71.43	**0.000143**	**0.000143**	**0.000143**	1	1	0.88391
	R10	87.75	11.85	88	11.9	87.5	12.55	72.5	**0.000258**	**0.000258**	**0.000258**	80	**0.000143**	**0.000143**	**0.000232**	1	1	0.86293
	R11	87.76	13.5	87.24	13.75	86.45	12.8	84.21	**0.00058**	**0.001099**	**0.028931**	92.11	**0.000143**	**0.000167**	**0.000392**	1	0.56473	0.72518
	R12	89.63	12.6	89.27	12.4	89.63	12.65	90.24	0.30762	**0.012228**	0.30762	85.37	**0.000143**	**0.000143**	**0.000162**	1	1	0.85212

Acc.: accuracy, N: number of channels, GA: genetic algorithm, BDE: binary differential evolution, BPSO: binary particle swarm optimization, KRU: Krusienski's 8-channel set, ALL: entire channel set.
Accuracies were averaged across the 20 repetitions of the algorithms for each subject.
P–values of the comparisons were obtained using a Wilcoxon signed-rank test. Then, the False Discovery Rate (FDR) was corrected using the Benjamini-Hochberg procedure. Significant p–values (< 0.05) are indicated in bold font.

An Asynchronous P300-Based Brain-Computer Interface Web Browser for Severely Disabled People

Víctor Martínez-Cagigal, Javier Gomez-Pilar, Daniel Álvarez, Roberto Hornero

Biomedical Engineering Group, E.T.S.I. Telecomunicación, University of Valladolid, Paseo de Belén 15, 47011, Valladolid, Spain

IEEE Transactions on Neural Systems and Rehabilitation Engineering
Volume 25, No. 8, August 2017, Pages 1332-1342

Abstract

This paper presents an electroencephalographic (EEG) P300-based braincomputer interface (BCI) Internet browser. The system uses the "odd-ball" row-col paradigm for generating the P300 evoked potentials on the scalp of the user, which are immediately processed and translated into web browser commands. There were previous approaches for controlling a BCI web browser. However, to the best of our knowledge, none of them was focused on an assistive context, failing to test their applications with a suitable number of end users. In addition, all of them were synchronous applications, where it was necessary to introduce a read-mode command in order to avoid a continuous command selection. Thus, the aim of this study is twofold: 1) to test our web browser with a population of multiple sclerosis (MS) patients in order to assess the use- fulness of our proposal to meet their daily communication needs; and 2) to overcome the aforementioned limitation by adding a threshold that discerns between control and non-control states, allowing the user to calmly read the web page without undesirable selections. The browser was tested with sixteen MS patients and five healthy volunteers. Both quantitative and qualitative metrics were obtained. MS participants reached an average accuracy of 84.14%, whereas 95.75% was achieved by control subjects. Results show that MS patients can successfully control the BCI web browser, improving their personal autonomy.

Keywords: Brain-computer interface (BCI), smartphones, asynchronous control, social networks, P300 event-related potentials, electroencephalography (EEG).

1. Introduction

The application of braincomputer interface (BCI) can improve the quality of life of those who have a disability that limits their ability to communicate, such as neurodegenerative diseases, traumatic brain injuries, Guillain Barré syndromes, degenerative muscle disorders, and other diseases that impair the neural pathways that control muscles or even the muscles themselves (Kübler and Birbaumer, 2008; Kübler et al., 2007; Wolpaw et al., 2000, 2002). BCI applications establish a communication system between the brain and the environment, translating the users intentions into device control commands. Even though there are a variety of methods for monitoring brain activity, electroencephalography (EEG) is commonly used due to its non-invasive nature. The electric potentials are recorded by means of placing several electrodes on the scalp (Wolpaw et al., 2000, 2002).

People who suffer multiple sclerosis (MS) are potential users of this kind of applications. MS is considered the most common autoimmune disorder that affects the central nervous system (World Health Organization, 2008). Twenty years after onset, up to 60% of the patients experience motor disability (World Health Organization, 2008). Although most people with MS have a normal or near-normal life expectancy, in rare cases, the disease can be terminal. MS is primarily an inflammatory disorder that leads to damage the myelin of brain and spinal cord nerve cells (Compston and Coles, 2008). This damage disrupts the ability of those neurons to communicate, resulting in a wide range of symptoms, including motor skill problems, cognitive deficit, or even psychiatric disorders (Compston and Coles, 2008).

MS patients could benefit from this technology for reducing their dependence. Due to its advance over the last few decades, Internet has caused a profound effect on peoples lives, becoming a global means of daily communication. However, web browsers are designed for healthy users, intended to be used with a keyboard and a mouse, but not with a small number of input signals (Mankoff et al., 2002). Therefore, it seems suitable to make the Internet accessible for those whose ability to communicate is restricted, in order to increase their autonomy, and thus their quality of life.

Email addresses: victor.martinez@gib.tel.uva.es (Víctor Martínez-Cagigal), javier.gomez@gib.tel.uva.es (Javier Gomez-Pilar), dalvgon@ribera.tel.uva.es (Daniel Álvarez), robhor@tel.uva.es (Roberto Hornero)

DOI: https://doi.org/10.1109/TNSRE.2016.2623381

There had been previously developed several attempts for controlling web browsers with BCI applications. The first ones used either slow cortical potentials (SCPs) or sensorimotor rhythms (SMR) as control mechanisms and were based on dichotomous approaches, using binary decision trees for selecting or rejecting commands (Karim et al., 2006; Bensch et al., 2007). Besides the slowness of the aforementioned approach, those browsers needed a supervisor who adjusted several parameters (e.g., reading speed, length of the reading pause, address book entries, etc.) (Karim et al., 2006; Bensch et al., 2007). In addition, both SCPs and SMR are endogenous signals, and it was necessary a long time so that the user learned how to control its own EEG activity (Karim et al., 2006; Bensch et al., 2007; Hinterberger et al., 2004). A few years later, Mugler et al. (2010) overcame the selection slow- ness of the dichotomous approach developing a BCI browser controlled via P300 evoked potentials based on the "oddball" paradigm (Farwell and Donchin, 1988). These potentials are produced in response to infrequent and particularly significant visual, auditory, or somatosensory stimuli about 300 ms after its elicitation (Wolpaw et al., 2002). Hence, training time was reduced because of their exogenous nature and the number of input signals drastically increased (Mugler et al., 2010; Sirvent Blasco et al., 2012). In addition, page links were tagged with an alphanumeric code and any link could be selected by entering the corresponding code with the selection matrix (Mugler et al., 2010). Sirvent Blasco et al. (2012) also used P300 evoked potentials as a control mechanism. However, instead of using the page tagging approach, one of the selection matrices was intended to work as a virtual mouse, whose commands allowed the user to move the cursor a variety of discrete pixel distances. Nevertheless, P300-based web browsers were synchronous processes and thus, it was needed to introduce several "read mode" commands for avoiding a continuous selection of items when the user wanted to calmly read the webpage, resulting in a rigid navigation (Mugler et al., 2010; Sirvent Blasco et al., 2012). For a truly free surfing, however, the synchronous mode is impractical because the system will deliver a selection even if the user is not paying attention to the stimulation (Pinegger et al., 2015). The latest BCI web browser was developed by Yu et al (Yu et al., 2012). The work was based on a two-dimensional BCI mouse that used SMR imagery and P300 potentials for controlling the horizontal and vertical movements, respectively. As stated above, its main limitation lied in the long required training time for learning to control the SMR activity.

The purpose of this study is twofold: 1) to design, develop and test a P300-based BCI web browser with a population of MS patients in order to assess the usefulness of our proposal to meet their daily communication needs; and 2) to provide the BCI web browser with an asynchronous approach in order to overcome the aforementioned limitations, by setting up a threshold which determines if the user is paying attention to the stimulation (control state) or, otherwise, is ignoring it (non-control state).

Table 1: Demographic and clinical characteristics of the participants

	User	Sex	Age	Motor disability	Cognitive ability	Sustained attention ability
MS	U01	30	F	Non-existent	Very high	Very high
	U02	31	M	Non-existent	High	Very high
	U03	43	M	Mild	Very high	High
	U04	47	F	Moderate	Normal	High
	U05	56	M	Moderate	Low	Very low
	U06	32	F	Non-existent	Normal	Normal
	U07	35	M	Non-existent	Very high	Very high
	U08	41	M	Non-existent	High	High
	U09	49	F	Non-existent	Normal	Very high
	U10	44	M	Mild	Normal	Low
	U11	41	F	Moderate	Normal	High
	U12	43	M	Moderate	Very high	Normal
	U13	44	M	Non-existent	High	High
	U14	52	M	Moderate	Very high	Normal
	U15	38	F	Non-existent	Normal	High
	U16	47	M	Moderate	Normal	Normal
CS	C01	23	M	-	-	-
	C02	31	M	-	-	-
	C03	23	M	-	-	-
	C04	31	M	-	-	-
	C05	22	M	-	-	-

CS: control subjects, MS: multiple sclerosis patients, F: female, and M: male.

2. Subjects and methods

2.1. Subjects

Sixteen MS patients (mean age 42.06 $\pm$ 7.47 years; 10 males, 6 females) and five healthy control subjects (CS) (mean age 26.00 $\pm$ 4.58 years; 5 males) were included in this study. MS participants were patients from the National Reference Centre on Disability and Dependence, located in León (Spain). The study was approved by the local ethics committee and all subjects gave their informed consent for participating in the study. Table 1 summarizes the demographic and clinical characteristics of all participants.

2.2. Description of the BCI Internet Browser

The application is composed of three different stages: data acquisition, EEG processing stage, and web surfing stage. As shown in Figure 1, data acquisition records the EEG signal and delivers it to the EEG processing phase. This stage controls the presentation of the stimuli and determines the selected command, which is delivered to the web surfing stage, responsible for interpreting the order and displaying the desired feedback.

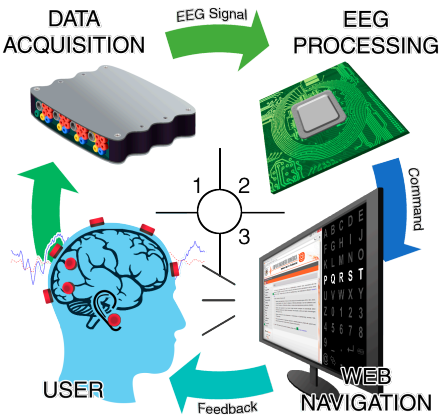


Figure 1: Structure of the BCI web browser. Three different stages compose the proposed system: data acquisition, EEG processing and web surfing.

2.2.1. Data Acquisition

The first stage records and pre-processes the EEG signals using a spatial and temporal filtering. Those signals were recorded using eight active electrodes placed on Fz, Cz, Pz, P3, P4, PO7, PO8 and Oz, according to the International 10–20 System distribution (Jasper, 1958), using a FPz electrode as a ground and referencing the system to the earlobe. This distribution is commonly used to record P300 potentials, mainly generated over the parietal cortex (Krusienski et al., 2006; Aloise et al., 2011; Corralejo et al., 2014). Electrodes were connected to a g.USBamp amplifier (g.Tec, Guger Technologies, Austria) with a sampling frequency of 256 Hz. Band-pass (0.1 Hz to 60 Hz) and notch (50 Hz power interference) filters were applied. In order to reduce the noise inside the recording, a common average reference (CAR) spatial filter was also applied. BCI2000 software (Schalk et al., 2004; Schalk and Mellinger, 2010) was used to control the presentation of the stimuli, record and save data on a laptop (Intel Core i7 2.40 GHz, 8 GB RAM, Windows 8.1). The selection matrices and the browser were displayed on an additional panoramic monitor (23-in screen) adjacent to the laptop.

2.2.2. EEG Processing Stage

The second stage, implemented in C++, processes the EEG signal received from the data acquisition. To this end, the system evokes the P300 potentials by means of the aforementioned "odd-ball" paradigm. In this paradigm, a target infrequent stimulus, which has to be attended, is presented among other more frequent background stimuli that have to be ignored (Wolpaw et al., 2002; Farwell and Donchin, 1988). When the user receives the target stimulus, a P300 evoked potential appears on the parietal cortex about 300 ms later. It has been widely documented that the amplitude of the P300 varies directly with the relevance of the eliciting events and inversely with the appearance probability of the target stimulus (Wolpaw et al., 2002; Farwell and Donchin, 1988). Specifically, we have used an application of the "odd-ball" paradigm known as row-col paradigm, whose stimuli are visual: matrix rows and columns are randomly flashed (Townsend et al., 2010). When the targets row or column are illuminated, a P300 potential is generated, used for figuring out what the desired command is (Wolpaw et al., 2002; Mugler et al., 2010; Farwell and Donchin, 1988; Sirvent Blasco et al., 2012; Yu et al., 2012). More specifically, each random stimulus lasts for 62.5 ms and then the screen remains unvarying for 125250 ms [19].

The application displays the browser on the left side of the screen and a selection row-col paradigm based matrix on the right side. Specifically, Google Chrome was selected as the target browser because it allows developers to comfortably program extensions (i.e., small software programs that can modify and enhance the functionalities of the browser). In order to provide a free and complete navigation, many com- mands are needed. Due to the large number of commands, the application uses alternatively two different matrices intended for different purposes (Figure 2)). "Navigation matrix" is the default one. Its small size (5×3) allows the user to quickly select the commands and thus, it is intended for web browsing. Hence, it contains navigation commands, such as scrolls, home page, reload, history forward and backward, among others. The other one is called the "keyboard matrix" (9×5), which is intended to write e-mails or fill out forms. For this reason, it contains all the alphanumeric characters and a variety of symbols commonly used on the Internet.

Once all the rows and columns have been flashed, it is needed to extract the most relevant features of the EEG signal. Because of the high sampling rate of the recordings relative to the low frequency of the P300 response, a dimensionality reduction for removing redundant features is beneficial for the real-time classification [23]. In this case, a subsampling of 20 Hz over an 800 ms window from the stimulus onset is applied [13], [19]. Therefore, each stimulus is considered a vector $\boldsymbol{f}$ of 128 features: 16 samples (20 Hz · 0.8 s) × 8 channels. As a result, the feature matrix of each character epoch would be $\boldsymbol{x} = [\boldsymbol{f}_1^T, \boldsymbol{f}_2^T, \ldots, \boldsymbol{f}_m^T]$, with $m = (N_r + N_c) \times N_s$ (sum of rows and columns×number of sequences).

The feature matrix is the input of the classification phase, which aims to determine the command the user wants to select. A linear classifier is used to determine whether there is a P300 potential in each stimulus or not. In this study, we used a step-wise linear discriminant analysis (SWLDA), a linear classifier that projects the data simultaneously minimizing the within-class covariance and maximizing the between-class covariance

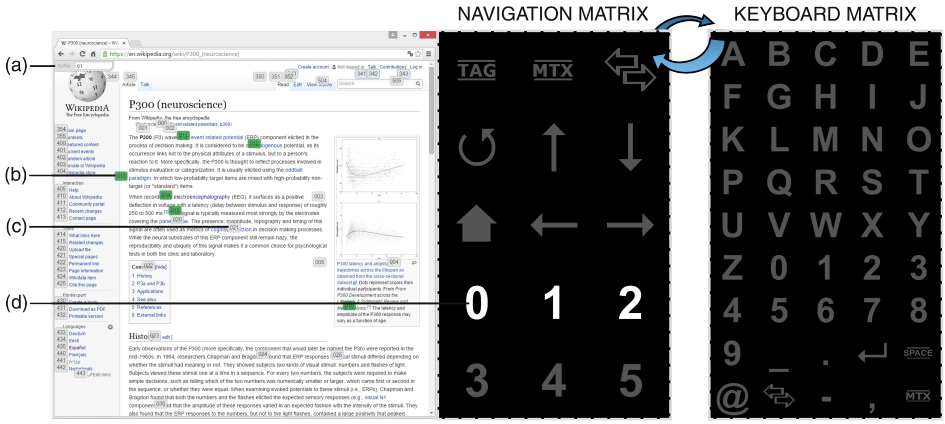


Figure 2: User application interface: current selection matrix is displayed on the right side of the screen and a Wikipedia web page on the left side. Selection matrix can be commuted by the user between navigation matrix (left) and keyboard matrix (right). As shown in the buffer (a), the user has previously selected "01" and thus, potential selections (b) are highlighted in green, while the rest of them (c) are colored in grey. In this shot, the fourth row of the current selection matrix is being illuminated (d).

(Keinosuke, 1990). In addition, the algorithm selects the most suitable features to be included in a multiple discriminant model, optimized for each user, reducing the dimensionality of the projection weight vector w. However, this solution and least-square regression are equivalent for binary classification tasks (Krusienski et al., 2008). The step-wise method decides to add or to remove a feature from the model by means of a combination of forward selection (add if pvalue < pin) and backward elimination (remove if pvalue > pout) steps, respectively (Krusienski et al., 2006, 2008). Therefore, significant differences (pvalue < 0.05) between models with and without the current evaluated feature are assessed to determine whether it provides discriminative information to the model or not. In this case, the discriminant function was restricted to contain a maximum of 60 features (Sirvent Blasco et al., 2012; Krusienski et al., 2006; Aloise et al., 2011; Corralejo et al., 2014; Krusienski et al., 2008), and the selection/elimination criteria was set up as pin = 0.10 and pout = 0.15, commonly applied in P300-based BCI studies (Krusienski et al., 2006; Corralejo et al., 2014; Krusienski et al., 2008; Sellers and Donchin, 2006; Sellers et al., 2006). Once the optimum weight vector w is computed, under the assumption that noise is normally distributed with equal covariances for both classes, the output of SWLDA is a log-likelihood ratio to belong to the positive class (i.e., presence of P300) (Narsky and Porter, 2013). This ratio is computed as the Euclidean distance between the projected data and the projected mean of the positive class, as follows:

$$\boldsymbol{l} = \|\langle \boldsymbol{w}, \boldsymbol{x} \rangle - \langle \boldsymbol{w}, \boldsymbol{\mu}_1 \rangle\|, \tag{1}$$

where $\boldsymbol{w}$ denotes the weight vector, $\boldsymbol{x}$ the feature matrix and $\boldsymbol{\mu}_1$ the mean of the positive class. In order to predict the selected item, it is necessary to turn the $\boldsymbol{l} \in \mathbb{R}^{m\times 1}$ vector into a matrix $\boldsymbol{P} \in \mathbb{R}^{N_r \times N_c}$ that indicates the probability of selecting each cell. Thus, for each matrix item p_i, the average of the log-likelihood scores of all the stimuli that belong to the same row and column is computed, as indicated in (2). Once the matrix $\boldsymbol{P}$ is calculated, the predicted item is the one that provides the maximum probability, $p_{char} = \max \boldsymbol{P}$.

$$p_i = \frac{1}{2N_s} \sum_{i=1}^{N_r N_c} l_{i \in \text{row} \cup \text{col}} \tag{2}$$

As stated above, row-col paradigm based selection matrices are synchronous processes. This means that the system will deliver a selection whether the user is paying attention to the stimulation (i.e., control state) or not (i.e., non-control state) Pinegger et al. (2015); Aloise et al. (2011). In this study, we have developed an asynchronous approach by placing a threshold (T) that is intended to distinguish between both states. When enough control and non-control state registers are recorded, the probability of the predicted item for each character epoch (p_{char}) is stored and labeled. In other words, two vectors are created by concatenating the predicted item probabilities for each character, which corresponds to control

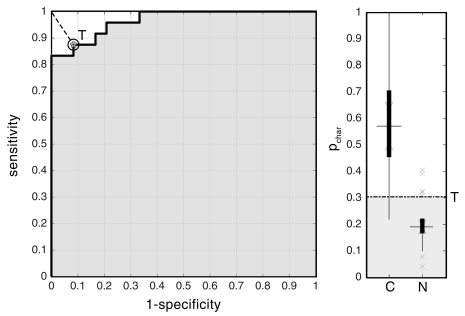


Figure 3: Threshold estimation for U07 user. (Left) ROC curve using different threshold values for the same subject data. Optimum threshold T (asterisk) was calculated as the point that maximizes the sensitivity and specificity pair (i.e., minimum Euclidean distance from (0,1) coordinates). (Right) Boxplots for control (C) and non-control (N) normalized probabilities of the predicted characters and users optimum threshold T (dash-dot black line).

and non-control selections. Due to the absence of attention, non-control probabilities are expected to be smaller than control ones and thus, it is expected that a constant threshold could discern between both states. Therefore, control and non-control vectors are fed as two different classes into a receiver operating characteristic (ROC) curve, a graphical plot that illustrates the performance of a binary classifier system as its discrimination threshold varies. The curve is created by plotting the true positive rate (i.e., sensibility) against the false positive rate (i.e., 1-specificity) for a set of threshold values. The custom threshold value for each user is chosen offline (i.e., before the evaluation sessions) as the point that fulfills the maximization of the sensitivity and specificity pair, looking for the best performance when distinguishing non-control and control states, as shown in Figure 3. Then, in online evaluation sessions, the probability of each new predicted item p_{char} is compared with the threshold value T: if $p_{char} > T$, the selection is classified as control state; if $p_{char} \leq T$, the selection is classified as non-control state. Finally, if the user intention is classified as a control state selection, it is delivered to the web surfing stage. Otherwise, the system considers it as a warning and asks the user for trying to select the command again.

2.2.3. Web Surfing Stage

The third and final stage was implemented in JavaScript as a Google Chrome extension. It is intended to receive and translate the user selections into browser commands and return a suitable feedback (Figure 2).

Firstly, the extension calculates how many nodes are on the current web page, where a node is any kind of clickable object, such as links, buttons or forms, among others. Then, those nodes are coded with the minimum number of digits using numbers from 0 to 5. As can be seen Figure 2, those numbers are included in the navigation matrix in order to increase the manageability of the application. Additionally, the "TAG" toggle controls the displaying of those codifications in form of tags allocated close to each link.

Thus, any link on the page can be executed by introducing its coding only using the navigation matrix (Mugler et al., 2010). Moreover, the application avoids the insertion of an additional "return" key to confirm the selection by automatically executing the link provided that the user has selected the needed number of characters, increasing the web surfing speed.

Feedback is provided to the user in several ways. On the one hand, when tag displaying is enabled, the extension initializes a buffer on the upper left corner of the screen that indicates what numbers were previously selected. In case of a selection error, user can remove the last selection with the left arrow command. On the other hand, potential selections (i.e., selections whose coding starts with the previously selected numbers) are highlighted in green, as shown in Figure 2.

2.3. Evaluation Procedure

During the assessment, all participants were sat down on a comfortable chair or on their own wheelchair, in front of a panoramic screen. Each user carried out four different sessions: two calibration sessions (Cal-I and Cal-II) and two evaluation sessions (Eval-I and Eval-II).

2.3.1. Calibration Sessions

The first calibration session was divided in two parts: classifier optimization and threshold calibration. Classifier optimization approximately lasted 24 min, and it was divided in 4 trials of 6 items (i.e., words composed of six characters). Users were asked for sequentially paying attention to those items while the matrix was flashing. Fifteen sequences were used (i.e., each row and column was illuminated 15 times in a single trial), so each desired character was highlighted 30 times. In order to keep the attention on the task, users were recommended to count how many times the target item was being flashed. In this initial session, only the keyboard matrix was displayed. Then, SWLDA was performed for assigning the optimum weights and number of sequences for each user. This custom classifier was used during the rest of the assessment.

Threshold calibration was composed of 8 trials with 6 items: half of those trials were intended to record the control state and the rest of them the non-control state. Navigation matrix was used in order to reduce the task time. Due to the variation on the optimum number of sequences for each user, the duration of this task differed between users (average trial duration, ME 3:12±1:08 min, CS 2:58±0:51 min). Control state was recorded with the

same procedure as the classifier optimization: asking the users for focusing their attention on specific series of commands. However, non-control state recordings followed a different procedure. A wide text was displayed on the left of the screen while navigation matrix was flashing. Users were asked to ignore the stimuli and read the text.

The second calibration session was only composed of another threshold calibration. That was necessary because the amplitude and latency of the P300 potentials usually vary between sessions, owing to the variation of the cap position on the scalp, user attention, attitude, among others (Picton, 1992). Therefore, recording the intensity of control and non-control state potentials in two different days increases the robustness of the asynchronous threshold customized for each user. The final threshold value was calculated as the average of both optimal thresholds.

Whether a user did not reach a minimum of 70% classification accuracy in the first calibration session, the classifier calibration was repeated in the second one. If after both sessions the user could not obtain more than 70% accuracy, considered as the minimum rate for experiment a satisfactory performance, the user was discarded of the assessment (Mugler et al., 2010; Corralejo et al., 2014). This case occurred three times in the MS subject group.

2.3.2. Evaluation sessions

Both evaluation sessions were intended to assess the quality of the web browser by means of setting different tasks. Nonetheless, threshold was not applied in the first one in order to determine if there is an improvement when it is applied (i.e., in the second one).

The first evaluation session was made up of five different tasks that required the use of the web browser. The four first ones were intended to assess the control state and the last one was only intended to assess the non-control state behavior. As pointed out earlier, tasks duration varied between users due to the optimal number of sequences for each one. However, a mean average time and its standard deviation are provided. The evaluation tasks were the following:

1) Link selection. Users had to scroll up and down a Wikipedia page and select one link (6 items, MS 4:01±1:31 min, CS 2:33±0:24 min).
2) Google searching. Users had to select the Google search form, write "BCI" inside it and select "ENTER" for running the search (8 items, MS 6:00±1:28 min, CS 4:28±1:03 min).
3) Publishing a tweet. Users had to select the Twitter form, write a two-character tweet and send it (6 items, MS 4:13±1:19 min, CS 2:38±0:31 min).
4) Writing an e-mail. Users had to read an inbox mail and reply it (13 items, MS 8:18±3:31 min, CS 6:18±2:13 min).
5) Passive reading. Users had to read a piece of news while ignoring the stimuli (10 items, MS 5:17±1:56 min, CS 4:17±0:45 min).

The second evaluation session was intended to assess the behavior of the web browser when threshold is enabled. It was made up of three slightly different tasks that involve the use of control and non-control states, alternating web page reading and web surfing:

1) Reading and link selection. Users had to scroll a Wikipedia page, read the information and select one link (8 items, MS 4:44±1:08 min, CS 4:18±1:44 min).
2) Publishing a tweet. Same procedure as Eval-I (6 items, MS 3:44±1:00 min, CS 3:25±1:28 min).
3) Active reading. Users had to read a piece of news, scrolling down the web page if needed (4 items, MS 2:20±0:55 min, CS 1:58±0:35 min).

The number of steps and the time needed to accomplish those tasks was recorded, as well as the mistakes and selections needed for solve them. With this information, a quantitative testing was performed, obtaining the users' accuracies and the false negative rate (FNR) for each task, defined as the ratio of false negatives (i.e., correct selection classified as non-state selection) to the total number of selections.

Also, a qualitative testing was performed in order to acquire a more accurate evaluation of the BCI web browser. At the end of the last session, users were asked for fulfilling a questionnaire. The survey consisted on 20 items to be ranked in a 7-point Likert scale that assessed the browser interface, its speed, the difficulty for selecting a command, the duration of sessions, users motivation, their expectations and their previous experience with BCI applications, among others. Additionally, one open-ended question allowed users to make personal suggestions for further improvement.

3. Results

3.1. Quantitative analysis

The results of the copy-spelling calibration sessions are presented in Table 2. The optimum number of sequences, the number of committed errors and calibration sessions accuracies for each user are shown. Accuracy is defined as the ratio of the number of correct selections in control state mode to the number of all performed selections, taking into account all the extra-selections made to correct the wrong ones. As can be seen, three MS patients could not obtain a minimum of 70% classifier accuracy and thus, they were removed from subsequent assessment. Also, as could be expected, CS users obtained a higher accuracy and a lower number of sequences than MS patients.

Tables 3 and 4 show the results of the evaluation tasks. For each task, the duration and the accuracy are presented. In addition, Table IV also indicates the FNR for each user and task. At the end of both tables, average session accuracy is shown for further comparison of the system behavior when threshold is disabled (Eval-I) or enabled (Eval-II).

Table 2: Copy-spelling calibration sessions results

	User	Cal-I		Cal-II		N_s
		Accuracy	WS(1)	Accuracy	WS(1)	
MS	U01	87.50%	3	79.17%	5	10
	U02	91.67%	2	87.50%	3	6
	U03	41.67%	14	75.00%	6	15
	U04	79.17%	5	95.83%	1	13
	U05	<70%	-	<70%	-	-
	U06	83.33%	4	66.67%	8	15
	U07	83.33%	4	91.67%	2	7
	U08	83.33%	4	70.83%	7	6
	U09	75.00%	6	95.83%	1	10
	U10	91.67%	2	75.00%	6	13
	U11	<70%	-	<70%	-	-
	U12	66.67%	8	70.83%	7	9
	U13	83.33%	4	66.67%	8	8
	U14	87.50%	3	87.50%	3	10
	U15	91.67%	2	75.00%	6	6
	U16	<70%	-	<70%	-	-
	Mean	**80.45%**	**5.14**	**79.81%**	**4.85**	**9.85**
	SD	13.65%	3.57	10.60%	2.54	3.29
CS	C01	100.0%	0	100.0%	0	7
	C02	100.0%	0	91.67%	2	11
	C03	100.0%	0	100.0%	0	6
	C04	100.0%	0	91.67%	2	10
	C05	100.0%	0	91.67%	2	9
	Mean	**100.0%**	**0**	**95.00%**	**1.2**	**8.60**
	SD	0.00%	0	4.56%	1.1	2.07

CS: control subjects, MS: multiple sclerosis patients, WS: wrong selections, N_s: optimal number of sequences for each subject.
(1) Each session was composed of a total number of 24 selections.
(2) Mean and SD were calculated regardless of the discarded users (i.e., those who could not reach a minimum of 70% accuracy in both calibration sessions).

3.2. Qualitative analysis

Satisfaction questionnaire results are shown in Table 5. It is noteworthy to mention that, in general, participants were quite satisfied with the BCI browser. All positive statements were rated above the mean value (4, neutral) and almost all negative ones were rated below it. As can be seen, the exceptions were the statements 3 and 13.

The third statement was intended to evaluate the speed of the browser. Some MS patients indicated that, in their opinion, it took much too long to surf the Internet with the BCI browser. In the thirteenth statement, both groups of users declared to be slightly happy that the assessment sessions were over.

Regarding the open-ended question, MS patients suggested increasing the command selection speed, adding a tab key command, trying to make the flashing less annoying, planning shorter sessions or trying to reduce the minimum number of sequences. CS users added that the number of symbols could be increased, for instance, by using other nested matrix. Also, they pointed out that sometimes they unintentionally focused their sight on adjacent cells.

4. Discussion

The application was assessed by sixteen MS patients and five CS users in four different days: two calibration and two evaluation sessions. Calibration sessions were intended to calculate the optimal SWLDA weights, number of sequences and threshold for each user, whereas evaluation sessions were intended to assess the BCI web browser completing different tasks. Moreover, a qualitative analysis was made giving a satisfaction questionnaire to the users.

As aforementioned above, three MS subjects were removed from the assessment due to their low classifier accuracy (<70%). This might be because their P300 evoked potentials were too attenuated and/or their latency was too variable. In addition, users could not be able to hold the attention while stimuli were presented. It is not surprising with regard to the U05 user owing to his lack of sustained attention capability, as shown in Table 1. However, the clinical characteristics of users U11 and U16 do not show an apparent reason for this behavior. Figure 4 shows two P300 potentials recorded in the Pz electrode for control and non-control states. It can be noticed that the P300 potentials of this kind of users were quite noisy and almost undetectable due to their low amplitude, which would explain the poor performance of their classifier.

CS users obtained higher calibration accuracy than MS patients in both sessions. Furthermore, their optimal number of sequences was lower than MS patients, so their browser surfing speed was higher. This is reflected in the questionnaire, specifically in the third statement, where CS users stated that it does not take too long to surf the Internet with the BCI browser, whereas MS patients requested a higher speed.

Regarding the first evaluation session, although all CS users could finish all tasks, it is worthy to note that not all MS patients were able to finish them. As expected due to its large number of minimal sequences, the fourth task, writing an e-mail, ended up as the most difficult one for both type of users, reaching the lower local accuracies (CS 90.96%, MS 71.11%). In contrast, link selection and publishing a tweet tasks were the easiest ones for CS and MS users, respectively. Average accuracies show that CS users got a better control of the browser (accuracy of 94.23%) than MS patients (accuracy of 77.46%). Nonetheless, even though some MS patients could not finish the tasks, five MS patients obtained average accuracies greater than 80%, and one of them (user U02) reached a perfect accuracy in all tasks (average accuracy of 100.00%). In the case of CS users, all of them obtained average accuracies greater than 80%, and two of them, C01 and C03, reached a perfect control of the browser (average accuracy of 100.00%).

In relation to the assessment of the threshold in Eval-I passive reading task, both type of users reached high accuracies (CS 96.00%, MS 85.77%). In fact, three CS users and three MS patients obtained a perfect accuracy, which

Table 3: Assessment results for the first evaluation session

	User[1]	Link selection		Google searching		*Tweet* publishing		E-mail writing		Passive reading		Average
		TIM	ACC	TIM	ACC	TIM	ACC	TIM	ACC	TIM	ACC	accuracy[2]
MS	U01	3:35	100.0%	6:05	100.0%	5:20	77.78%	15:04	66.67%	5:30	90.00%	79.54%
	U02	2:22	100.0%	4:02	100.0%	4:24	100.0%	5:17	100.0%	3:18	100.0%	100.0%
	U03	N.c.	25.00%	6:42	85.71%	N.c.	57.14%	N.c.	-	8:15	100.0%	61.11%
	U04	2:48	100.0%	4:45	100.0%	2:49	100.0%	6:24	92.86%	7:09	87.50%	96.97%
	U06	4:20	87.50%	8:05	72.73%	4:49	77.78%	8:54	75.00%	8:15	87.50%	77.27%
	U07	4:45	72.73%	N.c.	40.00%	3:37	83.33%	N.c.	50.00%	3:51	60.00%	63.33%
	U08	3:12	100.0%	6:20	100.0%	4:50	88.89%	6:58	100.0%	3:18	90.00%	97.20%
	U09	7:35	100.0%	N.c.	70.00%	3:20	100.0%	N.c.	62.50%	5:30	50.00%	82.35%
	U10	3:41	100.0%	N.c.	66.67%	5:01	75.00%	N.c.	62.50%	7:09	100.0%	76.00%
	U12	N.c.	66.67%	8:00	72.73%	7:02	63.64%	N.c.	53.33%	4:57	90.00%	62.79%
	U13	N.c.	71.43%	4:26	88.89%	2:13	100.0%	7:12	80.00%	2:45	90.00%	83.78%
	U14	3:48	75.00%	N.c.	75.00%	3:08	57.14%	N.c.	46.15%	5:30	80.00%	62.50%
	U15	N.c.	42.86%	5:31	70.00%	3:57	75.00%	N.c.	64.29%	3:18	90.00%	64.10%
	Mean	**4:01**	**80.09%**	**6:00**	**80.13%**	**4:13**	**81.21%**	**8:18**	**71.11%**	**5:17**	**85.77%**	**77.46%**
	SD	1:31	24.48%	1:28	17.84%	1:19	15.93%	3:31	18.67%	1:56	14.94%	14.24%
CS	C01	2:16	100.0%	3:40	100.0%	2:14	100.0%	3:39	100.0%	3:39	100.0%	100.0%
	C02	2:51	00.00%	5:56	75.00%	3:16	100.0%	9:06	75.00%	5:00	100.0%	82.50%
	C03	2:00	100.0%	3:16	100.0%	2:00	100.0%	4:20	100.0%	3:20	100.0%	100.0%
	C04	2:50	100.0%	4:45	100.0%	2:50	100.0%	6:06	92.31%	4:41	90.00%	96.97%
	C05	2:50	100.0%	4:44	100.0%	2:50	83.33%	8:18	87.50%	4:46	90.00%	91.67%
	Mean	**2:33**	**100.0%**	**4:28**	**95.00%**	**2:38**	**96.67%**	**6:18**	**90.96%**	**4:17**	**96.00%**	**94.23%**
	SD	0:24	0.00%	1:03	11.18%	0:31	7.45%	2:13	10.39%	0:45	5.48%	7.39%

TIM: task duration, ACC: task accuracy for each user, "n.c." (not completed) means that the user could not finish the task and thus, the task duration is not defined.
[1] MS patients U05, U11 and U16 were removed from the assessment because they could not obtain a minimum accuracy of 70% in calibration sessions.
[2] This average accuracy includes only the evaluation tasks that do not use the threshold (i.e., it does not include the passive reading results).

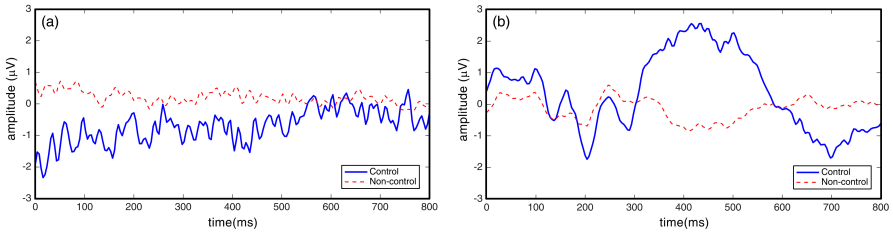


Figure 4: Average P300 potentials for control and non-control states recorded during the calibration sessions over the Pz electrode. User U05 (a) was discarded of the assessment due to his low classifier accuracy. His P300 potential is noisy and barely recognizable, in contrast to the user U07 (b), whose P300 potential has normal amplitude and latency.

a priori suggest that the use of threshold could improve the BCI browser performance.

In the second evaluation session, eight MS patients obtained average accuracies greater than 80%, and one of them (user U04) reached a perfect control of the browser. CS users achieved accuracies greater than 80% and 2 of them (users C03 and C05) obtained a perfect performance. Regarding the FNR, threshold causes an average of 4.61%±8.71% and 10.87%±14.33% of false negative errors to the total number of selections for CS and MS subjects, respectively. In addition, average accuracies of the second evaluation session (threshold enabled) are higher (CS 95.75%, MS 84.14%) than the obtained in the first one (CS 94.23%, MS 77.46%). However, it is noteworthy to mention that the improvement in terms of accuracy for CS subjects between both evaluation sessions is not significantly higher (Wilcoxon signedrank test, $p$value$=0.63$), probably because of their small number of overall errors. Therefore, the threshold avoided less uncertain selections for CS subjects than for MS ones. It suggests that, on subjects with full physical and cognitive capabilities, the introduction of the control state threshold

Table 4: Assessment results for the second evaluation session

User	Reading & Link			*Tweet* publishing			Active reading			Avg.
	TIM	FNR	ACC	TIM	FNR	ACC	TIM	FNR	ACC	ACC[1]
U01	6:05	0.30	92.31	N.c.	0.00	44.44	1:45	0.00	100.0	76.92
U02	4:02	0.25	91.67	2:31	0.13	85.71	2:48	0.00	100.0	92.00
U03	N.c.	0.13	100.0	N.c.	0.13	62.50	2:39	0.00	100.0	85.00
U04	3:50	0.00	100.0	2:48	0.00	100.0	1:53	0.00	100.0	100.0
U06	5:42	0.00	77.78	3:45	0.00	85.71	4:30	0.00	85.71	82.61
U07	N.c.	0.40	100.0	3:21	0.22	88.89	1:29	0.00	100.0	95.65
U08	3:22	0.00	70.00	3:42	0.09	81.82	2:08	0.00	75.00	76.00
U09	N.c.	0.17	83.33	N.c.	0.36	92.86	N.c.	0.00	75.00	86.67
U10	N.c.	0.08	63.64	N.c.	0.14	71.43	N.c.	0.00	75.00	68.18
U12	N.c.	0.17	33.33	N.c.	0.00	75.00	N.c.	0.33	83.33	72.00
U13	5:24	0.33	91.67	5:24	0.20	100.0	1:54	0.00	100.0	96.30
U14	N.c.	0.18	72.73	4:37	0.00	100.0	1:56	0.00	100.0	87.50
U15	N.c.	0.07	85.71	N.c.	0.56	66.67	N.c.	0.00	75.00	75.00
Mean	**4:44**	**0.16**	**81.71**	**3:44**	**0.14**	**81.16**	**2:20**	**0.03**	**83.93**	**84.14**
SD	1:08	0.13	18.78	1:00	0.17	16.69	0:55	0.09	11.77	10.08
C01	2:58	0.00	100.0	2:17	0.00	83.33	1:29	0.00	100.0	94.44
C02	7:02	0.23	92.31	5:20	0.10	100.0	2:44	0.00	60.00	89.29
C03	2:40	0.00	100.0	2:04	0.00	100.0	1:20	0.00	100.0	100.0
C04	4:32	0.00	100.0	4:36	0.25	87.50	2:20	0.00	100.0	95.00
C05	4:18	0.11	100.0	2:48	0.00	100.0	1:55	0.00	100.0	100.0
Mean	**4:18**	**0.07**	**98.46**	**3:25**	**0.07**	**94.17**	**1:58**	**0.00**	**92.00**	**95.75**
SD	1:44	0.10	3.44	1:28	0.11	8.12	0:35	0.00	17.89	4.48

MS patients indicated as Uxx, control subjects indicated as Cxx; TIM: task duration, ACC: accuracy for each user, AVG ACC: averaged accuracy, FNR: false negative rate (defined as the ratio of false negatives to the total number of selections), "N.c." (not completed) means that the user could not finish the task and thus, the task duration is not defined.

[1] MS patients U05, U11, and U16 were removed from the assessment because they could not obtain a minimum of 70% calibration accuracy.

does not provide an improvement in terms of accuracy, although it may provide a less demanding control. Regarding MS patients, there is an improvement of 6.68% between both sessions, although it does not provide a significant difference (Wilcoxon Signedrank Test, $p$value$=0.11$). Nonetheless, despite these advantages, a bad optimized threshold can lead to an increased required time to finish the proposed task due to false negative errors. This fact is clearly present in subject C02, who reached a perfect accuracy in the shared publishing a tweet task in both sessions. The required time for finishing the task was increased in the second one, since a 10% of the total selections were false negatives. This problem is caused by the inability of the threshold to follow the nonstationary changes of the EEG, which compromises a tradeoff between browser speed and selection accuracy. However, MS patients results reinforce the fact that the BCI browser performance is improved when the threshold is enabled, allowing end users to avoid selection mistakes when the intensity of their P300 potentials is not high enough for being considered as a strong deliberate selection.

Even though we are comparing the evaluation sessions in overall terms in order to assess the possible improvements when threshold is enabled, only one task (publishing a tweet) is strictly the same in both sessions. Due to the absence of threshold in the first session, none of the tasks involved the use of non-control state. Owing to the fact that the asynchrony management is one of the main contributions of the paper, we decided to slightly vary the first session tasks in order to involve changes between both states. For this reason, although the tasks are almost the same in both sessions, two of the second session tasks require commutations between the states, which increases the number of minimum selections and thus, the time to accomplish the tasks. Only one task remains unaltered, which is intended to be used for comparing the web browser performance when threshold is either enabled or disabled.

In fact, results from the shared task show a clear distinction between CS and MS users. In case of CS users, accuracy decreases from the first session (96.67%$\pm$7.45%) to the second one (94.17%$\pm$8.12%). As expected, the slightly higher amount of errors and the 7% rate of false negatives lead to an increase of the mean required time to accomplish the task (from 2:38$\pm$0:31 min to 3:25$\pm$1:28 min). Regarding the MS patients, accuracy remains practically the same (from 81.21%$\pm$15.93% to 81.16%$\pm$16.69%), although the required time decreases (from 4:13$\pm$1:19 min to 3:44$\pm$1:00 min). In addition, only a 14% of errors were false negatives. This behavior is caused by an increase of the number of users that could not finish the task, likely due to the intersession variability of the EEG, which cannot be followed by the constant threshold. Those nonstationary changes of the EEG are emphasized as more sessions are carried out without updating the custom classifier of each user and actually constitute one of the main limitations of the current BCI systems (Shenoy et al., 2006; Nicolas-Alonso et al., 2015).

Regarding the qualitative analysis results, as aforementioned, participants were quite satisfied with the BCI browser. However, thirteenth statement results show that the users were slightly happy that the assessment sessions were over. This fact reveals that its participation on the study implied an effort, which is an important aspect to take into consideration when the contents and duration of the sessions are planned. Nevertheless, it is worth to note that users were willing to participate in further studies. As previously stated, for MS patients, the item with the lowest rating was the speed. However, browser speed is directly related to classifier accuracy, which is calculated in the calibration sessions. A more robust classifier, either using a more sophisticated training algorithms or having more training samples, could obtain higher calibration accuracy. Thus, it could reduce the optimal number of sequences in order to experience a faster navigation. This issue does not appear in CS users, probably because their number of sequences was lower and they committed fewer mistakes than MS patients. In contrast, the top-rated aspect was the application interface, due to its simplicity and user-friendliness.

Users also pointed out that, sometimes, they unintentionally focused their attention on adjacent cells. This is-

Table 5: Questionnaire results for the post-study assessment of the BCI web browser

	Statement	CS		MS	
		Mean	SD	Mean	SD
1	I am not used to surf the Internet	1.2	0.4	3.2	2.3
2	I have found interesting to use the BCI web browser	6.4	0.5	6.2	0.9
3	In my opinion, it takes much too long to surf the Internet with the BCI browser	4.0	1.0	4.8	0.9
4	My expectations for the browser were completely met	6.4	0.5	5.7	1.2
5	I was bored during the assessment sessions	2.8	1.3	2.7	2.0
6	I can imagine myself using this browser in my daily life	3.4	1.1	4.6	1.7
7	I became impatient during the assessment sessions	2.2	0.8	2.8	1.8
8	I found the assessment sessions entertaining	5.2	0.4	5.4	1.2
9	It was stressful to concentrate when it was required	2.0	1.0	2.9	1.7
10	I would gladly carry out more testing sessions with the BCI browser	6.6	0.5	5.6	1.2
11	The assessment sessions made me feel tired	2.8	1.8	3.0	1.9
12	User interface is intuitive and easy to understand	5.4	2.1	5.9	1.3
13	I am happy that the assessment sessions are over	4.6	0.9	4.2	1.9
14	I found it easy to select keyboard matrix commands	4.6	1.7	5.3	1.3
15	I found it difficult to select navigation matrix commands	2.4	1.1	2.5	1.3
16	I like computers and information technologies	6.0	1.0	4.9	2.4
17	Flickering stimuli are annoying	1.8	0.8	3.7	2.1
18	I would love to participate in other similar studies	6.0	1.0	5.2	1.5
19	The duration of the assessment sessions was too long	3.0	1.0	2.4	1.4

CS: control subjects, MS: multiple sclerosis patients.
Each statement was rated on a 7-point Likert scale where 0 means "strongly disagree", 4 means "neutral" and 7 means "strongly agree".

sue is known as the adjacency-distraction problem and is inherent in the row-col paradigm (Townsend et al., 2010). As its name suggests, adjacent non- target flashes distract users and cause selection errors, which are commonly found in adjacent cells or in those that belong to the targets row and column. Specifically, the percentage of this kind of selection errors out of the total number of errors are 100% (out of 14 errors) and 87.57% (out of 141 errors) for CS and MS users, respectively. The probabilities of randomly selecting one of those cells in the navigation and keyboard matrices are 44.66%–66.66% and 28.88%–35.55% (depending on the position of the target cell), respectively. Therefore, it is clear that most errors are due to this problem. A possible solution to the adjacency-distraction problem is to use the checkerboard paradigm, which solves both this and the double-flash problem (Townsend et al., 2010).

Table **??** shows a comparison between previous BCI browsers and our present study. Besides the fact that P300 evoked potentials and node tagging makes the proposed browser faster and more self-sufficient than other previous approaches (Karim et al., 2006; Bensch et al., 2007), the main advantage is the asynchrony management. In this study, a control state threshold was implemented, avoiding the use of a constant supervision or a rigid "read command" (Karim et al., 2006; Bensch et al., 2007; Mugler et al., 2010; Sirvent Blasco et al., 2012). This approach allows users to calmly experiment a free surfing navigation while the system continuously detects the users intentions based on their attention. It is noteworthy to mention that strictly speaking, due to the nature of the "odd-ball" paradigm, the application actually remains synchronous. The use of this control state threshold removes undesired selections when the user is ignoring the stimuli, but it does not avoid synchronous selections, owing to the fact that the matrices keep flashing. Nevertheless, it is common to classify applications that use control state detection strategies as "asynchronous BCIs", which is a widely used term in the BCI literature (Pinegger et al., 2015; Aloise et al., 2011; Zhang et al., 2008; Panicker et al., 2011; Pan et al., 2013). This strategy makes the control smoother and probably less demanding, which is an advantage for end users who are suffering from physical limitations. In addition, our approach was tested with a bigger patient database than previous studies (Karim et al., 2006; Bensch et al., 2007; Mugler et al., 2010), whereas our CS subject pool is limited in comparison with (Mugler et al., 2010) and (Pinegger et al., 2015). However, CS subjects are not potential users of this kind of applications and thus, their results cannot be generalized to any disease.

Regarding the web browser performance, the differences in the accuracies obtained by CS users and MS patients suggest that the reason lies on the symptoms of the disease. For the MS patients, it has been observed a highly variable classifier performance during the sessions. Nevertheless, previous studies stated that P300-based BCI systems can be controlled by severely disabled people, regardless of their degree of disability (Kübler and Birbaumer, 2008; Kübler et al., 2007; Corralejo et al., 2014; Picton, 1992). As can be seen, notwithstanding its lower performance compared to CS subjects, MS patients average accuracy (84.14%) is higher than the ones reported in the previous approaches tested by ALS patients (Karim et al., 2006; Mugler et al., 2010; Sirvent Blasco et al., 2012). Significant differences are found between our work and the accuracies provided by Mugler et al (Mugler et al., 2010) when using a Mann-Whitney U–

Table 6: Comparison between previous BCI browsers and present study

Author	Control signal	Type of signal	Functionalities		Assessment		
			Link selection	Asynchrony	Subjects		Accuracy
Karim et al. (2006)	SCP	Endogenous	Dichotomous tree	Supervision	1	ALS	80.00%
Bensch et al. (2007)	SCP/SMR	Endogenous	Dichotomous tree	Supervision	4 2	ALS CS	- -
Mugler et al. (2010)	P300	Exogenous	Node tagging	Read command	3 10	ALS CS	72.00% 93.40%
Sirvent Blasco et al. (2012)	P300	Exogenous	Virtual cursor	Read command	4	CS	93.00%
Yu et al. (2012)	P300/SMR	Both	Virtual cursor	Not needed	7	CS	93.21%
Present study	P300	Exogenous	Node tagging	Threshold	16 5	MS CS	84.14% 95.75%

Test ($p$value$= 0.0193$). However, although care must be taken owing to the differences between both diseases, accuracies show a higher overall performance for the disabled subject population in comparison with other previous studies. Additionally, the cognitive disability that commonly appears in MS patients is rarely presented on ALS patients, since their neurologic damage is generally focused on motor neurons. For this reason, it is suggested to test the BCI web browser with ALS patients as a future line of research in order to get a better comparison between both diseases. Similarly, CS users average accuracy (95.75%) is also slightly higher than those reported in previous approaches (Mugler et al., 2010; Sirvent Blasco et al., 2012; Yu et al., 2012), although it is not significantly different (Mann-Whitney U–Test, p-value> 0.05). These results reveal that the use of a threshold for discerning between control and non-control states could be a useful contribution for further asynchronous BCI P300-based systems. In addition, a control state threshold appears to be a more comfortable solution for users than a "read mode" command, because it eliminates the need for being attentive to select a command when the user wants to rest or to read a web page. Even though these results show that our BCI web browser has successfully allowed severely disabled people to experiment a truly free Internet surfing, we can point out some limitations. As previously indicated, the major drawback of this kind of applications is the classifier performance variability between sessions and users. Reducing this variability and increasing the classification accuracy by using more suitable processing techniques in both feature extraction and selection could improve the robustness of the system (Shenoy et al., 2006; Nicolas-Alonso et al., 2015). In addition, control state threshold is calculated directly over the SWLDA scores and thus, it depends on the classifier performance of each user. Using algorithms that are not dependent on the classifier, such as residual steady-state visually evoked potentials, could improve the application performance (Pinegger et al., 2015; Aloise et al., 2011; Zhang et al., 2008). Another limitation is the impossibility to alternate between lower case and capital letters, essential for fulfilling user and password forms, and it should be a further improvement. To conclude, although we have included a significant amount of symbols, an additional nested matrix with extra symbols (e.g., tab key, ampersand, slashes, brackets, etc.) would contribute to access any web page in the address bar.

5. Conclusion

An asynchronous P300-based BCI web browser has been designed, developed and evaluated. The system processes the EEG signal of the users, and P300 potentials are elicited using a visual "odd-ball" row-col paradigm composed of two different matrices, which contains navigation and keyboard commands. Those commands are sent to a Google Chrome extension, which traduces them and gives visual feedback to the users. The browser has been tested with five CS users and sixteen MS patients. Results show that our BCI web browser can successfully meet their daily communication needs, allowing end users to surf the Internet in an intuitive way. In addition, the average accuracies achieved by CS and MS users (95.75% and 84.14%, respectively) are higher than that reported in previous approaches. In fact, significant differences have been found (p-value> 0.0193) between our results and the accuracies reported in previous studies for disabled subjects. However, care must be taken owing to the fact that end users suffered from different diseases. Therefore, control state threshold appears to be an appropriate solution for developing further asynchronous BCI systems.

Acknowledgments

This work was partially supported by the projects TEC2014-53196-R from 'Ministerio de Economía y Competitividad' (MINECO) and FEDER and VA037U16 from 'Consejería de Educación de la Junta de Castilla y León'. V. Martínez-Cagigal was in receipt of a 'Promoción de Empleo Joven e Implantación de la Garantía Juvenil en I+D+i' grant from MINECO and the University of Valladolid.

References

Aloise, F., Schettini, F., Aricò, P., Leotta, F., Salinari, S., Mattia, D., Babiloni, F., Cincotti, F., 2011. P300-based braincomputer interface for environmental control: an asynchronous approach. Journal of Neural Engineering 8 (2), 025025.

Bensch, M., Karim, A. a., Mellinger, J., Hinterberger, T., Tangermann, M., Bogdan, M., Rosenstiel, W., Birbaumer, N., 2007. Nessi: An EEG-Controlled Web Browser for Severely Paralyzed Patients. Computational intelligence and neuroscience 2007, 71863.

Compston, A., Coles, A., 2008. Multiple sclerosis. Lancet 372 (9648), 1502–1517.

Corralejo, R., Nicolás-Alonso, L. F., Álvarez, D., Hornero, R., 2014. A P300-based braincomputer interface aimed at operating electronic devices at home for severely disabled people. Medical & Biological Engineering & Computing 52 (10), 861–872.

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

Hinterberger, T., Schmidt, S., Neumann, N., Mellinger, J., Blankertz, B., Curio, G., Birbaumer, N., 2004. Brain-computer communication and slow cortical potentials. IEEE Transactions on Biomedical Engineering 51 (6), 1011–1018.

Jasper, H. H., 1958. The ten-twenty electrode system of the international federation. Electroencephalography and Clinical Neurophysiology 10, 371–375.

Karim, A. A., Hinterberger, T., Richter, J., Mellinger, J., Neumann, N., Flor, H., Kubler, A., Birbaumer, N., 2006. Neural Internet: Web Surfing with Brain Potentials for the Completely Paralyzed. Neurorehabilitation and Neural Repair 20 (4), 508–515.

Keinosuke, F., 1990. Introduction to statistical pattern recognition. Academic Press Inc.

Krusienski, D., Sellers, E., McFarland, D., Vaughan, T., Wolpaw, J., 2008. Toward enhanced P300 speller performance. Journal of Neuroscience Methods 167 (1), 15–21.

Krusienski, D. J., Sellers, E. W., Cabestaing, F., Bayoudh, S., McFarland, D. J., Vaughan, T. M., Wolpaw, J. R., 2006. A comparison of classification techniques for the P300 Speller. Journal of neural engineering 3 (4), 299–305.

Kübler, A., Birbaumer, N., 2008. Brain-computer interfaces and communication in paralysis: Extinction of goal directed thinking in completely paralysed patients? Clin. Neurophysiol. 119 (11), 2658–2666.

Kübler, A., Nijboer, F., Birbaumer, N., 2007. Brain-Computer Interfaces for communication and motor control – perspectives on clinical application. In: Toward Brain-Computer Interfacing, 1st Edition. MA: The MIT Press, pp. 373–391.

Mankoff, J., Dey, A., Batra, U., Moore, M., 2002. Web accessibility for low bandwidth input. Proceedings of the fifth international ACM conference on Assistive technologies (9), 17–24.
URL http://dl.acm.org/citation.cfm?id=638255

Mugler, E. M., Ruf, C. a., Halder, S., Bensch, M., Kübler, A., 2010. Design and implementation of a P300-based brain-computer interface for controlling an internet browser. IEEE Transactions on Neural Systems and Rehabilitation Engineering 18 (6), 599–609.

Narsky, I., Porter, F. C., 2013. Statistical analysis techniques in particle physics: fits, density estimation and supervised learning. John Wiley & Sons.

Nicolas-Alonso, L. F., Corralejo, R., Gomez-Pilar, J., Álvarez, D., Hornero, R., 2015. Adaptive semi-supervised classification to reduce intersession non-stationarity in multiclass motor imagery-based braincomputer interfaces. Neurocomputing 159, 186–196.
URL http://linkinghub.elsevier.com/retrieve/pii/S0925231215001496

Pan, J., Li, Y., Zhang, R., Gu, Z., Li, F., 2013. Discrimination between control and idle states in asynchronous SSVEP-based brain switches: A pseudo-key-based approach. IEEE Transactions on Neural Systems and Rehabilitation Engineering 21 (3), 435–443.

Panicker, R. C., Puthusserypady, S., Sun, Y., 2011. An asynchronous P300 BCI with SSVEP-based control state detection. IEEE Transactions on Biomedical Engineering 58 (6), 1781–1788.

Picton, T. W., 1992. The P300 wave of the human event-related potential. Journal of Clinical Neurophysiology 9 (4), 456–479.

Pinegger, A., Faller, J., Halder, S., Wriessnegger, S. C., Müller-Putz, G. R., 2015. Control or non-control state: that is the question! An asynchronous visual P300-based BCI approach. Journal of Neural Engineering 12 (1), 014001.

Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N., Wolpaw, J. R., 2004. BCI2000: A general-purpose brain-computer interface (BCI) system. IEEE Transactions on Biomedical Engineering 51 (6), 1034–1043.

Schalk, G., Mellinger, J., 2010. A Practical Guide to BrainComputer Interfacing with BCI2000, 1st Edition. Springer, London.

Sellers, E. W., Donchin, E., 2006. A P300-based brain-computer interface: initial tests by ALS patients. Clinical Neurophysiology 117 (3), 538–548.

Sellers, E. W., Krusienski, D. J., McFarland, D. J., Vaughan, T. M., Wolpaw, J. R., 2006. A P300 event-related potential brain-computer interface (BCI): the effects of matrix size and inter stimulus interval on performance. Biological psychology 73 (3), 242–52.
URL http://www.ncbi.nlm.nih.gov/pubmed/16860920

Shenoy, P., Krauledat, M., Blankertz, B., Rao, R. P. N., Müller, K.-R., 2006. Towards adaptive classification for BCI. Journal of neural engineering 3 (1), R13–R23.

Sirvent Blasco, J., Iáñez, E., Úbeda, A., Azorín, J., 2012. Visual evoked potential-based brainmachine interface applications to assist disabled people. Expert Systems with Applications 39 (9), 7908–7918.

Tomori, O., Moore, M., 2003. The neurally controllable internet browser (BrainBrowser). CHI '03 extended abstracts on Human factors in computing systems - CHI '03, 796.
URL http://portal.acm.org/citation.cfm?doid=765891.765997

Townsend, G., LaPallo, B. K., Boulay, C. B., Krusienski, D. J., Frye, G. E., Hauser, C. K., Schwartz, N. E., Vaughan, T. M., Wolpaw, J. R., Sellers, E. W., 2010. A novel P300-based brain-computer interface stimulus presentation paradigm: moving beyond rows and columns. Clinical Neurophysiology 121 (7), 1109–1120.

Wolpaw, J. R., Birbaumer, N., Heetderks, W. J., McFarland, D. J., Peckham, P. H., Schalk, G., Donchin, E., Quatrano, L. A., Robinson, C. J., Vaughan, T. M., 2000. Brain-computer interface technology: a review of the first international meeting. IEEE Transactions on Rehabilitation Engineering 8 (2), 164–173.

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., Vaughan, T. M., 2002. Brain-computer interfaces for communication and control. Clinical Neurophysiology 113 (6), 767–91.

World Health Organization, 2008. Atlas: multiple sclerosis resources in the world. Vasa.
URL http://medcontent.metapress.com/index/A65RM03P4874243N.pdf

Yu, T., Li, Y., Long, J., Gu, Z., 2012. Surfing the internet with a BCI mouse. Journal of Neural Engineering 9 (3), 036012.

Zhang, H., Guan, C., Wang, C., 2008. Asynchronous P300-based brain-computer interfaces: a computational approach with statistical models. IEEE Transactions on Biomedical Engineering 55 (6), 1754–63.

Towards an accessible use of smartphone-based social networks through brain-computer interfaces

Víctor Martínez-Cagigal, Eduardo Santamaría-Vázquez, Javier Gomez-Pilar, Roberto Hornero

Biomedical Engineering Group, E.T.S.I. Telecomunicación, University of Valladolid, Paseo de Belén 15, 47011, Valladolid, Spain

Expert Systems with Applications
Volume 120, April 2019, Pages 155-166

Abstract

This study presents an asynchronous P300-based Brain–Computer Interface (BCI) system for controlling social networking features of a smartphone. There are very few BCI studies based on these mobile devices and, to the best of our knowledge, none of them supports networking applications or are focused on an assistive context, failing to test their systems with motor-disabled users. Therefore, the aim of the present study is twofold: (i) to design and develop an asynchronous P300-based BCI system that allows users to control Twitter and Telegram in an Android device; and (ii) to test the usefulness of the developed system with a motor-disabled population in order to meet their daily communication needs. Row-col paradigm (RCP) is used in order to elicitate the P300 potentials in the scalp of the user, which are immediately processed for decoding the user's intentions. The expert system integrates a decision-making stage that analyzes the attention of the user in real-time, providing a comprehensive and asynchronous control. These intentions are then translated into application commands and sent via Bluetooth to the mobile device, which interprets them and provides visual feedback to the user. During the assessment, both qualitative and quantitative metrics were obtained, and a comparison among other state-of-the-art studies was performed as well. The system was tested with 10 healthy control subjects and 18 motor-disabled subjects, reaching average online accuracies of 92.3% and 80.6%, respectively. Results suggest that the system allows users to successfully control two socializing features of a smartphone, bridging the accessibility gap in these trending devices. Our proposal could become a useful tool within households, rehabilitation centers or even companies, opening up new ways to support the integration of motor-disabled people, and making an impact in their quality of life by improving personal autonomy and self-dependence.

Keywords: Brain-computer interface (BCI), smartphones, asynchronous control, social networks, P300 event-related potentials, electroencephalography (EEG).

1. Introduction

Brain–Computer Interfaces (BCI) are able to establish a communication system between our brains and the environment, making it possible to control devices with our brain signals. Such bypassing requires the monitoring of brain activity, which is commonly accomplished using electroencephalography (EEG) due to its portability, non-invasiveness, and low-cost (Wolpaw et al., 2000). Hence, electric potentials are recorded by placing a set of electrodes over the user's scalp (Wolpaw et al., 2000, 2002). The main motivation of BCI systems has always been to improve the quality of life of motor-disabled people, which usually contributes to reduce the accessibility gap in different fields. Thus, end users can take advantage of this novel technology to reduce their dependence, regardless of their disability. These motor-disabilities could have been caused by neurodegenerative diseases, traumas, muscle disorders, or any illness that impair the neural pathways that control muscles or the muscles themselves (Wolpaw et al., 2000, 2002; Kübler et al., 2007; Kübler and Birbaumer, 2008). Moreover, BCI systems may use a wide variety of control signals to detect the user's intentions in real time (Wolpaw et al., 2002). In particular, exogenous signals, such as P300 evoked potentials, are commonly used to assure the efficacy of the systems with any motor-disabled user. These potentials are produced in response to infrequent and particularly significant stimuli about 300 ms after their onset (Wolpaw et al., 2002).

The rapid growth of the Internet in the last decades has caused a huge impact on people's lives, bringing entirely new ways of everyday communication. This impact has been enlarged by the popularity of the smartphones, which provide a continuous Internet connection. In fact,

Email addresses: victor.martinez@gib.tel.uva.es (Víctor Martínez-Cagigal), eduardo.santamaria@gib.tel.uva.es (Eduardo Santamaría-Vázquez), javier.gomez@gib.tel.uva.es (Javier Gomez-Pilar), robhor@tel.uva.es (Roberto Hornero)

DOI: http://doi.org/10.1016/j.eswa.2018.11.026

it is estimated that there are 4.9 billion of unique mobile users in the world, reaching a market penetration of 66% (Kemp, 2017). Their functionalities cover from managing finances to reading news, including watching videos, shopping, playing games or searching for information. However, it is worthy to note that more than the 56% of the time spent is dedicated to socializing (i.e., social media and instant messaging), both in everyday and working environments (Ipsos MORI and Google, 2017). Currently, there are 2.8 billion of active social media users, and 91.4% of them access social media with their smartphones or tablets (Ipsos MORI and Google, 2017). Despite this development, the accessibility of these devices is still restricted for motor-disabled people that are unable to use accurately their hands and fingers.

Motor disabilities comprise the limitations on people's physical functioning that hinder their full and effective interaction with the environment and society (World Health Organization, 2011). These impairments may be caused by: (i) neurodegenerative diseases, such as multiple sclerosis, amyotrophic lateral sclerosis, Friedreich's ataxias, etc.; (ii) congenital conditions, such as cerebral palsy, polymalformative syndromes, myotonic dystrophies, etc.; or (iii) traumas, such as strokes or spinal cord injuries, among others. It is estimated that the world average prevalence rate of disability for adult people is 15.6%, which ranges from 11.8% in higher income countries to 18.0% in lower ones (World Health Organization, 2011). Moreover, diseases and traumas are not the only cause that can lead to develop a motor disability, but also the natural ageing contributes in a high extent. In fact, older people are disproportionately represented in disability populations and thus, everybody is susceptible to develop a motor disability at some point in their lives (World Health Organization, 2011). In this respect, BCI applications represent a novel technology from which disabled people can benefit to reduce their dependence.

From an expert and intelligent systems point of view, BCIs utilize artificial intelligent techniques to replace, restore, enhance or supplement the natural central nervous system outputs of their users (Hill and Wolpaw, 2016). To this end, BCIs should comprise a decision-making stage that interprets neural activity and determines users' intentions or emotions. Moreover, several BCIs include an adaptive engine that learns from the experience, modifying classifier weights and features while the user controls the system (Atkinson and Campos, 2016). These systems can be trained to react to changes in the EEG signals that could reflect: (i) emotions (Blondet et al., 2013; Atkinson and Campos, 2016), (ii) road drowsiness (Da Silveira et al., 2016), (iii) driving stress (Chen et al., 2017), (iv) mental effort (Zammouri et al., 2018), (v) attention (Aloise et al., 2011; Pinegger et al., 2015; Martínez-Cagigal et al., 2017a), (vi) motor imagery (Wolpaw et al., 2002), or (vii) event-related responses (Luck, 2014), among others. Accordingly, BCIs play a potential role as knowledge-based systems in many clinical and industrial applications.

In recent years, some studies have attempted to apply BCI systems to mobile devices with the aim of controlling a wheelchair (Jayabhavani et al., 2013), robots (Ma et al., 2015), or detecting the user's emotions (Blondet et al., 2013). Despite the popularity of the smartphones and tablets these days, there are very few studies in the scientific literature that aim to control any of their functionalities. These studies are limited to dial numbers in cell phones (Wang et al., 2011; Chi et al., 2012), accept incoming calls (Katona et al., 2014), call contacts (Campbell et al., 2010; Wang et al., 2011), spell words (Obeidat et al., 2017; Elsawy et al., 2017), or play a simple racing game (Wu et al., 2014). Possibly the work of Elsawy and Eldawlatly (2015) is the one that relates more closely to the topic, which allows users to open pre-installed apps and visualize the image gallery. Nevertheless, to the best of our knowledge, none of those studies has been focused on providing a high-level control of a smartphone or tablet, nor making social apps accessible to disabled people. Furthermore, it is well known that disabled users generally reach lower accuracies than healthy users (Wolpaw et al., 2002; Sellers and Donchin, 2006; Martínez-Cagigal et al., 2017a) and thus, the assessment of BCI systems with end users is essential for ensuring a fair evaluation. Since these studies have not been tested with a disabled population, their reliability may be compromised in real life situations.

The purpose of this study is twofold: (i) to design and develop a practical BCI-based application that allows disabled people to access social media with any smartphone or tablet; and (ii) to evaluate it with a population of motor-disabled people in order to assess the usefulness of our proposal to meet their daily communication needs. With the objective of providing a comprehensive social networking support, we consider that the system should implement both a social network and an instant messaging applications. In this case, the application will provide a complete control of Twitter and Telegram, which currently have more than 317 and 100 millions of mobile active users, respectively (Kemp, 2017). Moreover, the application will monitor users' attention and apply a dynamic asynchronous control management (Martínez-Cagigal et al., 2017a). As a result, the expert system will only deliver conscious selections, eliminating the need of read-mode commands or external supervisors.

2. Subjects

Eighteen motor-disabled subjects (MDS, mean age: 47.63 $\pm$ 9.53 years; 11 males, 8 females) and ten healthy control subjects (CS, mean age: 26.10 $\pm$ 3.45 years; 8 males, 2 females) were included in this study. MDS participants were recruited from the National Reference Centre on Disability and Dependence, located in León (Spain). All subjects gave their informed written consent to participate in the study, previously approved by the local ethics

Table 1: Demographic and clinical data of the participants

	User	Sex	Age	DD	Disease
Motor-Disabled subjects	**M01**	F	48	90%	Stroke
	M02	M	46	80%	Spinal cord injury
	M03	F	38	93%	Friedreich's ataxia
	M04	M	39	98%	Spinal cord injury
	M05	F	49	78%	Friedreich's ataxia
	M06	M	31	76%	Cerebral palsy
	M07	M	52	99%	Cerebral palsy
	M08	M	44	90%	Friedreich's ataxia
	M09	M	47	69%	Cerebral palsy
	M10	M	67	87%	Cerebral palsy
	M11	M	62	86%	Myotonic dystrophy
	M12	M	47	90%	Polymalformative syndrome
	M13	F	66	94%	Friedreich's ataxia
	M14	F	40	88%	Friedreich's ataxia
	M15	M	38	98%	Spinal cord injury
	M16	M	50	80%	Spinal cord injury
	M17	F	42	89%	Cerebral palsy
	M18	F	45	84%	Spinal cord injury
Control subjects	**C01**	M	25	0%	-
	C02	M	25	0%	-
	C03	M	24	0%	-
	C04	M	25	0%	-
	C05	M	25	0%	-
	C06	M	32	0%	-
	C07	M	24	0%	-
	C08	M	25	0%	-
	C09	F	23	0%	-
	C10	F	33	0%	-

F: female, M: male, DD: degree of disability.

committee. Table 1 summarizes the clinical and demographic characteristics of all participants. As can be noticed, all MDS present moderate or high degrees of motor disability (mean: 86.42% $\pm$ 8.58%), caused by different diseases: stroke (1), spinal cord injuries (5), Friedreich's ataxias (5), cerebral palsies (5), polymalformative syndrome (1), and myotonic dystrophy (1).

3. Methods

As shown in Fig. 1, the developed BCI application involves three main entities, which communicate among themselves: (i) the user, which involves the EEG signal acquisition; (ii) the laptop, which generates the visual stimuli, processes the signal, decodes the user's intentions and translates them into commands; and (iii) the mobile device, which interprets those commands and provides visual feedback to the user. The methodology that is applied to each stage, as well as the evaluation procedure, are described below.

3.1. Signal acquisition

EEG signals from users were recorded using eight active electrodes, placed on Fz, Cz, Pz, P3, P4, PO7, PO8 and Oz, according to the International 10–20 System distribution (Jasper, 1958). The system was referenced to the earlobe,

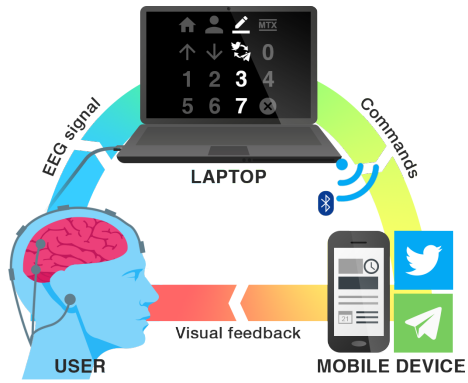


Figure 1: Structure of the BCI social network application. The EEG signal of the user is sent to the laptop, which processes it, decodes the user's intentions and translates them into commands in real time. These commands are finally sent to the device (i.e., smartphone or tablet) via Bluetooth, which interprets them and provides visual feedback to the user.

using the Fpz electrode as a ground. Electrodes were connected to a g.USBamp amplifier (g.Tec, *Guger Technologies*, Austria) with a sampling frequency of 256 Hz. As a pre-processing stage, band-pass (0.1–60 Hz), notch (50 Hz) and common average reference (CAR) filters were applied. BCI2000 platform was used to record the data, display and process the stimuli (Schalk et al., 2004).

3.2. Signal processing

The exogenous nature of P300 evoked potentials avoids training (Wolpaw et al., 2002). Furthermore, the number of different commands that can be selected by the user is extremely large whether the *odd-ball* paradigm is used (Farwell and Donchin, 1988; Wolpaw et al., 2002; Martínez-Cagigal et al., 2017a). In this paradigm, an infrequent target stimulus, which has to be attended, is presented among other distracting background stimuli that have to be ignored. When the user attends to the target stimulus, a P300 potential appears mainly on the parietal and occipital cortex (Farwell and Donchin, 1988; Wolpaw et al., 2002; Martínez-Cagigal et al., 2017a). We used an extension of the *odd-ball* paradigm, known as row-col paradigm (RCP), for decoding the users' intentions (Townsend et al., 2010). In the RCP, a matrix containing the commands that control the BCI application is displayed, whose rows and columns are randomly flashed. The user, who has to stare at the desired command, will generate a P300 potential when the target's row or column is illuminated (Farwell and Donchin, 1988; Wolpaw et al., 2002; Townsend et al., 2010; Martínez-Cagigal

et al., 2017a; Martínez-Cagigal and Hornero, 2017; Obeidat et al., 2017).

Social media apps in general and, particularly, Twitter and Telegram, have some key functionalities that should be controlled. In this regard, owing to the fact that not only the RCP matrices have to include control commands, but also alphanumeric characters and symbols, our application uses alternatively two different matrices: (i) main matrix, and (ii) keyboard matrix (see Fig. 2). The first one is intended to control the main functionalities of Twitter and Telegram, such as loading the home view, opening a new *tweet* or chat, visualizing a profile or contact, toggling between both social networks or scrolling the current view. The second one, by contrast, is intended to write texts and fill out forms. Both matrices can be freely toggled between themselves if the user selects the command "MTX".

Due to the high sampling rate of the EEG recordings relative to the low frequency of the P300 potential response, a dimensionality reduction is beneficial for the real-time classification (Krusienski et al., 2008). In order to extract the most relevant features of the EEG signal, a sub-sampling of 20 Hz is applied on the first 800 ms from the stimulus onset (i.e., 16 samples per stimulus and channel). Then, channels are concatenated, returning a vector of 128 features per stimulus (Corralejo et al., 2014; Martínez-Cagigal et al., 2017a). Afterwards, the extracted feature vectors of each stimulus are processed by a linear classifier, which determines the presence (i.e., positive class) or the absence (i.e., negative class) of a P300 evoked potential. Step-wise linear discriminant analysis (SWLDA) was used in this study, with $p_{in} = 0.10$ and $p_{out} = 0.15$ as selection/elimination criteria and a maximum of 60 selected features for each input vector (Krusienski et al., 2006, 2008; Corralejo et al., 2014; Martínez-Cagigal et al., 2017a; Martínez-Cagigal and Hornero, 2017). Even though SWLDA has a simple implementation, it delivers similar performances and lower computational cost in comparison with more complex methods, which makes it a popular algorithm for the P300 classification problem (Krusienski et al., 2006, 2008; Blankertz et al., 2011; Zhang et al., 2016; Martínez-Cagigal et al., 2017b). This method calculates a projection of the input data that simultaneously minimizes the within-class and maximizes the between-class covariances (Keinosuke, 1990). Thus, the probability score of finding a P300 in the i-th illumination is computed using the Euclidean distance between the projected data and the projected mean of the positive class (Narsky and Porter, 2013), as follows:

$$l_i = 1 - \|\langle \mathbf{w}, \mathbf{x}_i \rangle - \langle \mathbf{w}, \mu_i \rangle\| \tag{1}$$

where $\mathbf{w}$ is the weight vector, computed in a calibration session; $\mathbf{x}_i$ denotes the feature vector, and μ_i the mean of the positive class. The probability of selecting a certain command j is computed as the average of the scores of all the stimuli that belong to its row and column, as indicated in (2). Therefore, the output selected command is the one that provides the maximum average probability (i.e., $p_s = \max \mathbf{p}$) (Martínez-Cagigal et al., 2017a).

$$p_j = \frac{1}{N} \sum l_{i \in row \cup col} \tag{2}$$

RCP-based matrices are synchronous processes, which means that the system will deliver a selection even if the user is not paying attention to the visual stimulation (Aloise et al., 2011; Pinegger et al., 2015; Martínez-Cagigal et al., 2017a). This fact severely restricts the autonomy of the application, needing an external supervisor or implementing a read-mode command that could pause the system for a fixed number of seconds. In our application, we have implemented a dynamic asynchronous control management by monitoring the user's attention (Martínez-Cagigal et al., 2017a). The method works as follows: (i) EEG signals of the user paying attention (i.e., control state) and ignoring (i.e., non-control state) the stimuli are recorded in a calibration session; (ii) the signals are processed and the final selected command probabilities p_s are stored in both control and non-control arrays; (iii) the arrays are fed into a receiver operating characteristic (ROC) curve for determining the optimum threshold that maximizes the sensitivity-specificity pair; (iv) the custom threshold value T for each user is then used online. In the online sessions, the selected command probability is compared with the threshold in real-time. If $p_s > T$, the selection is accepted and the command is sent via Bluetooth to the mobile device; otherwise, the selection is rejected and the system encourages the user to try to select the command again.

3.3. Application

It has been recently reported that 98.8% of the smartphones that are sold these days either use Android or iOS (International Data Corporation, 2017). In fact, Android has an 83.4% of the worldwide smartphone market share, while iOS has a 15.4% (International Data Corporation, 2017). For this reason, and taking into account that Android is a free open platform, we have developed our application for this operating system. Whether the application is used for the first time, the user is asked to login the Twitter account and to register the telephone number to Telegram. Switching between both services is also handled by a toggle command that can be selected by the user. Fig. 3 shows several snapshots of the final application, whose main functionalities are described below.

Twitter. Defined as a popular free social networking service that allows users to broadcast public small messages (up to 280 characters), known as *tweets*. Although it was originally developed as an online service, its mobile activity reaches more than 317 million of active users, which makes Twitter one of the most installed social networking services in smartphones or tablets nowadays (Kemp,

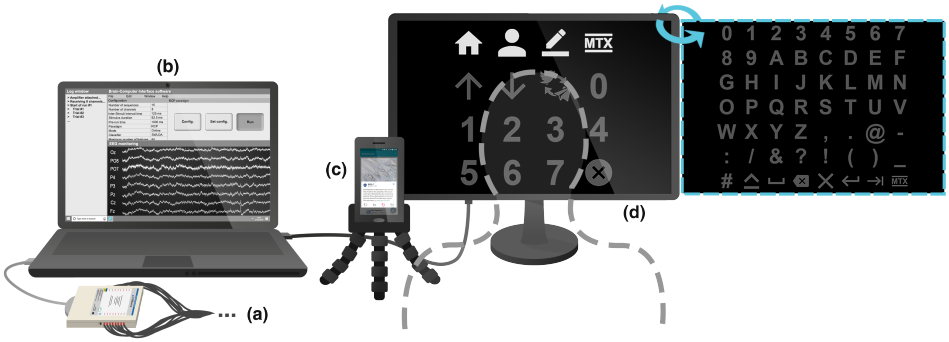


Figure 2: Evaluation setup from the point of view of the user: (a) EEG acquisition unit, (b) laptop that monitors the EEG signal, processes it and generates the stimuli; (c) smartphone on a small tripod, close enough to the user for receiving the visual feedback; (d) panoramic screen that displays the stimuli. Both matrices are depicted: (left) main matrix, whose first row is currently flashed; and (right) keyboard matrix, which can be toggled by the user through the "MTX" command.

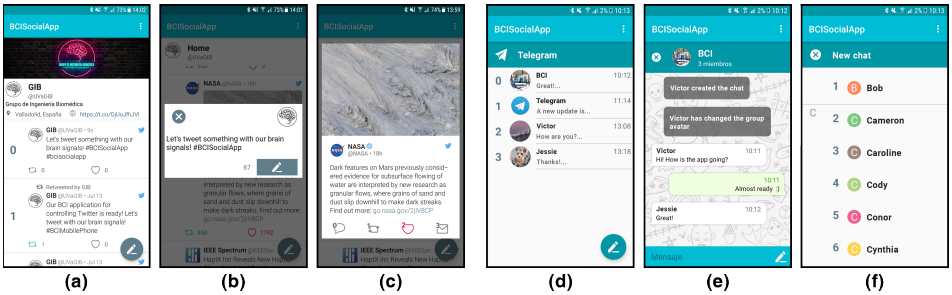


Figure 3: Snapshots of the BCI social networking application: (a) Twitter's profile timeline, (b) dialog for writing *tweets*, (c) *tweet* view, (d) Telegram's conversation list, (e) Telegram's group, and (f) contact list.

2017). Our BCI application implements the entire set of Twitter functionalities, including both the possibility of interacting with: (i) "tweets", writing, answering, "retweeting", or making them as favorite; and (ii) accounts, surfing among profiles, or sending direct messages.

Telegram. Defined as a non-profit cloud-based instant messaging service that allows users to send encrypted messages and exchange files of any type in real-time. Even though it has a desktop version, its popularity is extended thanks to the mobile application, which has more than 100 million of active users and has become the most popular instant messaging app in several countries (Kemp, 2017). Our BCI application covers its main functionalities, including the possibility of interacting with individual chats, groups and channels through real-time messages, or creating new chats with any contact that is stored in the device.

3.4. Evaluation procedure

The evaluation setup is depicted in Fig. 2. During the assessment, participants were sat on a comfortable chair or on their own wheelchair, in front of a panoramic screen, as well as in front of a smartphone on a small tripod. The screen was connected to a laptop (Intel Core i7 @ 2.6 GHz, 16 GB RAM, Windows 10), which executed the signal processing stage and sent the commands to the mobile device (Samsung Galaxy S7, 4GB RAM, Android 7.0) via Bluetooth. The assessment was composed by three different sessions: the first two intended to calibrate the system, and the last one intended to evaluate the BCI application.

Calibration 1. The first session was intended to compute the optimal parameters for each user, such as the

number of sequences (i.e., repetitions of the stimuli), the classifier's weight vector, and the asynchronous threshold value. Firstly, users were asked to pay attention to 6 items in 4 different trials (i.e., to spell 4 words composed of 6 characters). Due to its larger size, the keyboard matrix was used and the number of sequences was fixed in 15. During this calibration, users were encouraged to count how many times the target character was being flashed, in order to keep attention to the task. After these runs, SWLDA was trained, returning the weight vector and the most appropriate number of sequences for each user. The latter is computed as the minimal number of repetitions that reaches a 100% of accuracy using the training data. Hereinafter, the trained SWLDA model and the optimal number of sequences for each user were used in the online sessions. Note that training data was composed of 5400 observations per subject (6 items $\times$ 4 trials $\times$ 15 seq. $\times$ [7 rows + 8 columns]). Then, the first stage of threshold calibration was performed. Composed of 8 trials with 6 items, the calibration was intended to record signals of both control and non-control states. Thus, users were asked to pay attention to 4 trials, and to ignore the flashings of the remaining 4 (e.g., by reading a text).

Calibration 2. The second session was intended to finish the threshold calibration for increasing the overall performance. The objective was to record additional data in order to create a most robust asynchronous threshold that could be adapted to the inter-session variability of the participants (Picton, 1992; Martínez-Cagigal et al., 2017a). Hence, users were asked to spell 4 trials and ignore 4 trials more, all of them composed by 6 items. It is noteworthy that both stages of the threshold calibration were performed using the main matrix, aiming to reduce the task time due to its smaller size. Then, thresholds for both sessions were calculated as the optimal points of the ROC curves using control and non-control classes. Finally, the optimal threshold value was computed as the average of them.

Evaluation. The third session was intended to assess the performance and the quality of the developed BCI system. The evaluation session, strictly online, was made up of 6 different tasks, whose difficulty increased progressively. It is worthy to mention that the duration of each task varied among users due to their different optimal number of sequences. These tasks are described below, together with the ideal number of selections and the matrices that are required to finish them.

i) Toggling between Twitter and Telegram. Using Twitter, users had to scroll down and up the timeline and toggle to Telegram (3 items, main matrix).
ii) Retweeting a *tweet*. Using Twitter, users had to scroll down the timeline, select one *tweet* and retweet it (4 items, main matrix).
iii) Writing a new *tweet*. Using Twitter, users had to open the form to write a new *tweet* and spell "hello" (7 items, both matrices).
iv) Checking the profile and answering a *tweet*. Using Twitter, users had to visit their own profile, select the last *tweet* and answer "great!" (11 items, both matrices).
v) Creating a new chat. Using Telegram, users had to select one contact, create a new chat, and spell "how are you?" (11 items, both matrices).
vi) Chatting with someone. Using Telegram, users had to select one chat from the conversations list, in which the interlocutor had said: "hi! how are you?", and reply with "fine, and you?" (12 items, both matrices).

During the evaluation session, both quantitative and qualitative metrics have been registered. With regard to the quantitative measures, the number of correct selections, errors, sequences and the time that it takes to accomplish each task have been noted down. As a result, accuracies and output characters per minute (OCM) for each task have been calculated. Accuracy is defined as the percentage of correct selections to the total number of selections. It is worthy to note that the selections that have not overcome the asynchronous threshold have not been considered errors, since they have not been sent to the final device. OCM, calculated by dividing the total number of selections by the duration of the task, is an online metric that estimates the true communication rate of the system (Speier et al., 2013). Although information transfer rate (ITR) has traditionally been used in this respect, several authors pointed out that ITR makes assumptions that are usually incorrect in online BCI systems (Speier et al., 2013; Yuan et al., 2013). ITR assumes that: (i) all possible selections are equally probable, (ii) the system is memoryless, and (iii) a synchronous paradigm is used. In online systems where users are allowed to correct selection errors, ITR may return counterintuitive results when two different users type the same word and one shows lower speed, but returns a higher ITR. Since correcting an error implies to successfully spell two or more commands, the ITR increases because the decrease in accuracy weighs less than the increase in extra selections. Moreover, ITR requires the number of possible selections (i.e., n), as well as the reached accuracy. Despite that the latter is a global metric, n varies if more than one RCP matrix is used, hindering the generalization of ITR values. In addition, ITR assumes that commands are sequentially selected following a constant speed, without pauses. Therefore, the estimation is biased in asynchronous-based BCI systems. It is also noteworthy that the ITR estimation is incorrect if the subject did not perform any error, returning an infinite value. According to this rationale, ITR is replaced by OCM considering the nature of the proposed BCI system.

Regarding the qualitative testing, users were asked to fulfill a questionnaire at the end of the session. The survey was composed of 20 items that had to be ranked in a 7-point Likert scale (Likert, 1932). These items assessed

the subjective opinions of the users in regard to the application speed, interface, accessibility, the duration of the sessions, the users' motivation and their expectations, among others. Moreover, an additional open-ended question was included to collect their personal suggestions for further improvements. It is noteworthy that optimal number of sequences and trained SWLDA models, previously computed in the calibration sessions for each subject, were used thereinafter in the online evaluation session.

4. Results

Results of the calibration sessions are depicted in Table 2, where training accuracies, optimal number of sequences, and percentage of error selections in control-state recordings are detailed for each user. As can be noticed, 4 MDS could not obtain training accuracies higher than 70%. Since 70% is usually considered as the minimal acceptable accuracy in the BCI literature, they were discarded from the subsequent assessment (Kübler et al., 2001; Kleih et al., 2010; Corralejo et al., 2014; Martínez-Cagigal et al., 2017a). Quantitative results of the evaluation sessions are shown in the Table 3, including the duration, the final accuracy and the OCM of each task. Moreover, their averages and the number of sequences of each user are also detailed. Questionnaire results are finally depicted in Table 4, which specifies the statements and the ranks provided by the users. Values range from 1 (i.e., totally disagree), to 7 (i.e., totally agree), where 4 means a neutral response. Note that positive and negative statements are alternated in order to reduce the acquiescence bias (Likert, 1932). With regard to the final open-ended question, two users demanded to get rid of the conductive gel, and one user demanded more speed.

Table 2: Calibration sessions results

User	Classifier		Threshold	
	TA	N_s	A1	A2
M01	67.0%	15	-	-
M02	89.0%	10	41.7%	83.3%
M03	92.0%	14	50.0%	50.0%
M04	100%	9	95.8%	95.8%
M05	100%	7	95.8%	70.8%
M06	100%	7	83.3%	77.8%
M07	8.0%	15	-	-
M08	100%	10	87.5%	68.2%
M09	100%	13	100%	72.2%
M10	100%	13	79.2%	79.2%
M11	57.0%	15	-	-
M12	100%	12	83.3%	87.5%
M13	56.0%	15	-	-
M14	100%	9	66.7%	58.3%
M15	100%	13	83.3%	87.5%
M16	100%	14	95.8%	87.5%
M17	89.0%	15	50.0%	33.3%
M18	100%	7	95.8%	91.7%
C01	100%	11	100%	91.7%
C02	100%	6	100%	97.2%
C03	100%	13	95.8%	95.8%
C04	100%	7	100%	95.8%
C05	100%	5	87.5%	91.7%
C06	100%	8	91.7%	91.7%
C07	100%	8	95.8%	100%
C08	100%	4	77.8%	91.7%
C09	100%	8	100%	100%
C10	100%	7	100%	95.8%

The prefix "M" stands for motor-disabled subjects, whereas "C" indicates the control subjects; "TA" stands for training accuracy; N_s indicates the number of sequences of each user; and "A1" and "A2" indicate the accuracy in the first and second threshold sessions, respectively.

5. Discussion

Four MDS were discarded from the assessment due to their low training accuracy (<70%) (Kübler et al., 2001; Kleih et al., 2010; Corralejo et al., 2014; Martínez-Cagigal et al., 2017a), probably because their P300 potentials were too attenuated or their latencies were too variable (Table 2). Since there are subjects with the same diseases that do not show this effect, the rationale behind it lies in indirect problems related to attention capability or gaze control. In particular, M01 exhibited lack of sustained attention capability; M07 suffered from essential tremors; M11 was unable to open his eyes properly; and M13 reported nystagmus, which causes involuntary eye movements, resulting in limited vision and lack of control over gaze. Fig. 4 depicts two sample ERPs recorded over channels Pz and Cz, one from M16, who could finish all tasks; and the other one from M07, who was discarded from the assessment. In contrast to the response of M16, the P300 potential of M07 is quite noisy and unrecognizable, which would explain the poor performance of his classifier in the training stage.

Unsurprisingly, quantitative results of the evaluation session (Table 3) show that CS obtained higher overall accuracies (92.3% $\pm$ 6.3%) than MDS (80.6% $\pm$ 12.9%). In fact, this difference in performance was demonstrated to be significant (Wilcoxon Signed-rank Test, p-value $=$ 0.0375). Furthermore, the required number of sequences for CS was significantly lower (Wilcoxon Signed-rank Test, p-value $=$ 0.0155) than for MDS, which used 7.7 $\pm$ 2.7 and 10.93 $\pm$ 2.84 sequences, respectively. Consequently, the bits per minute rate for CS (2.06 $\pm$ 0.73) was also higher than for MDS (1.47 $\pm$ 0.40), producing also significant differences (Wilcoxon Signed-rank Test, p-value $=$ 0.0498). The less number of sequences, the higher output bits per minute. This assures a faster navigation through the application and thus, CS took less time than MDS to finish the tasks. These findings reinforce the necessity of assessing the reliability of BCI systems with end users.

With regard to the complexity of these tasks, the average durations of the Table 3 show a clear increase as the users advance through the tasks, especially for CS. However, the average accuracies for each of them does not show a constant decreasing, which could be expected at

Table 3: Evaluation session results

User	Task 1			Task 2			Task 3			Task 4			Task 5			Task 6			N_s	Avg. acc,	Avg. OCM
	Dur.	Acc.	OCM	Dur.	Acc.	OCM	Dur.	Acc.	OCM	Dur.	Acc.	OCM	Dur.	Acc.	OCM	Dur.	Acc.	OCM			
M02	01:52	66.7%	1.61	04:55	60.0%	2.04	06:09	66.7%	1.46	06:09	63.6%	1.79	08:59	63.6%	1.22	01:02	100%	1.94	10	65.2%	1.58
M03	03:06	100%	1.29	04:42	57.1%	1.49	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	14	72.7%	1.41
M04	01:05	100%	2.76	02:29	100%	2.42	04:36	100%	1.52	06:32	100%	1.68	09:12	77.8%	0.98	03:11	100%	1.57	9	95.1%	1.51
M05	01:05	100%	2.76	01:27	100%	2.76	03:35	85.7%	1.95	05:05	90.9%	2.16	04:31	100%	1.99	05:39	100%	1.94	7	95.6%	2.11
M06	01:33	100%	1.94	03:37	85.7%	1.94	03:04	100%	2.28	04:40	100%	2.36	05:31	100%	2.17	05:50	84.6%	2.23	7	94.3%	2.18
M08	01:33	100%	1.94	02:04	100%	1.94	05:07	85.7%	1.37	08:18	58.3%	1.45	03:29	40.0%	1.44	04:49	71.4%	1.45	10	71.1%	1.50
M09	02:01	100%	1.49	03:22	100%	1.49	06:39	100%	1.05	10:07	81.8%	1.09	05:35	50.0%	1.07	n.c.	n.c.	n.c.	13	84.4%	1.15
M10	02:01	66.7%	1.49	03:22	40.0%	1.49	07:43	75.0%	1.04	09:20	63.6%	1.18	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	13	63.0%	1.20
M12	01:52	66.7%	1.61	03:43	100%	1.61	07:07	75.0%	1.12	09:20	81.8%	1.18	11:02	60.0%	0.91	n.c.	n.c.	n.c.	12	76.3%	1.15
M14	01:05	66.7%	2.76	02:16	100%	1.76	05:38	85.7%	1.24	09:08	58.3%	1.31	06:26	66.7%	1.86	05:05	60.0%	1.97	9	68.8%	1.62
M15	02:01	100%	1.49	04:02	66.7%	1.49	07:02	87.5%	1.14	10:07	72.7%	1.09	11:58	100%	1.00	10:30	100%	1.05	13	88.2%	1.12
M16	02:10	66.7%	1.38	02:54	100%	1.38	07:75	75.0%	1.01	10:54	90.9%	1.01	12:53	91.7%	0.93	11:19	100%	0.97	14	89.8%	1.02
M17	02:20	100%	1.29	04:39	83.3%	1.29	10:08	66.7%	0.89	11:40	45.5%	0.94	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	15	65.5%	1.01
M18	01:05	100%	2.76	01:27	100%	2.76	03:35	100%	1.95	05:27	100%	2.02	06:26	100%	1.86	06:48	92.3%	1.91	7	98.0%	2.02
Mean	**01:46**	**88.1%**	**1.90**	**03:13**	**85.2%**	**1.85**	**06:01**	**84.8%**	**1.39**	**08:13**	**77.5%**	**1.48**	**07:31**	**77.2%**	**1.40**	**07:49**	**89.8%**	**1.67**	**10.93**	**80.6%**	**1.47**
SD	00:35	16.6%	0.60	01:09	20.7%	0.49	02:02	12.5%	0.43	02:22	18.4%	0.47	03:10	22.4%	0.48	03:15	14.9%	0.44	2.84	12.9%	0.40
C01	01:42	100%	1.76	02:16	100%	1.76	05:38	100%	1.24	07:47	90.9%	1.41	08:05	90.9%	1.30	09:12	91.7%	1.36	11	93.8%	1.38
C02	00:56	100%	3.23	01:14	100%	3.23	03:04	85.7%	2.28	04:40	100%	2.36	04:51	100%	2.17	05:31	100%	2.27	6	97.9%	2.37
C03	02:01	100%	1.49	04:02	83.3%	1.49	07:43	85.7%	0.908	10:07	100%	1.09	10:30	100%	1.00	13:01	92.3%	1.14	13	94.2%	1.10
C04	01:05	100%	2.76	02:10	66.7%	2.77	03:35	100%	1.95	05:27	81.8%	2.02	05:39	100%	1.55	09:43	73.3%	2.12	7	85.2%	1.95
C05	00:47	100%	3.87	01:02	100%	3.87	02:33	100%	2.74	03:54	90.9%	2.83	04:03	100%	2.61	04:36	100%	2.97	5	98.0%	2.90
C06	00:56	100%	4.30	01:14	100%	3.23	03:04	71.4%	2.28	06:14	100%	1.12	08:05	100%	1.30	09:12	66.7%	1.36	8	86.7%	1.57
C07	01:14	100%	2.42	02:04	60.0%	2.42	03:35	57.1%	1.95	05:27	81.8%	2.02	06:53	91.7%	1.50	07:22	81.8%	1.74	8	79.6%	1.84
C08	00:37	100%	4.84	00:50	100%	4.84	02:03	100%	3.42	03:07	100%	3.53	03:14	90.9%	3.26	03:41	91.7%	3.40	4	95.8%	3.55
C09	01:14	100%	2.42	01:39	100%	2.42	04:06	100%	1.71	06:38	91.7%	1.81	05:39	100%	1.86	06:26	100%	1.94	8	98.0%	1.90
C10	01:05	100%	2.76	01:49	80.0%	2.77	03:35	100%	1.95	05:27	100%	2.02	05:27	90.9%	1.86	06:26	91.7%	2.02	7	93.9%	2.06
Mean	**01:10**	**100%**	**2.99**	**01:50**	**89.0%**	**2.88**	**03:54**	**90.0%**	**2.04**	**05:53**	**93.7%**	**2.02**	**06:15**	**96.4%**	**1.84**	**07:31**	**89.0%**	**2.03**	**7.7**	**92.3%**	**2.06**
SD	00:25	0.0%	1.08	00:55	15.6%	0.98	01:39	15.1%	0.71	02:00	7.5%	0.76	02:10	4.6%	0.68	02:48	11.5%	0.72	2.7	6.3%	0.73

The prefix "M" stands for motor-disabled subjects, whereas "C" indicates the control subjects; "Dur." indicates the task duration; "Acc." indicates the task accuracy for each user; "OCM" stands for Output Characters per Minute; N_s indicates the number of sequences of each user; and "n.c." stands for "not completed", which means that the user could not finish the task and thus, durations, accuracies and OCM are not defined. Note that users M01, M07, M11 and M13 were discarded from the assessment because they could not obtain a minimum accuracy of 70% in the calibration sessions.

Table 4: Questionnaire results

No.	Statement	MDS		CS	
		Mean	SD	Mean	SD
1	I found interesting to use the BCI social networking application	**6.07**	1.07	**6.00**	0.94
2	I found it difficult to control the system	**2.86**	1.79	**2.70**	1.34
3	My expectations for the application were completely met	**5.29**	1.64	**5.90**	0.99
4	I was bored during the assessment sessions	**2.14**	1.56	**3.50**	1.96
5	I found the assessment sessions entertaining	**5.57**	1.65	**4.80**	1.40
6	I can imagine myself using this BCI application in my daily life	**4.29**	2.34	**2.60**	1.84
7	It was stressful to concentrate when it was required	**3.00**	1.75	**2.60**	1.71
8	The application works smoothly	**4.71**	1.44	**5.80**	1.03
9	The duration of the calibration sessions was too long	**2.43**	1.74	**3.70**	1.89
10	User interface is intuitive and easy to understand	**4.79**	1.76	**5.70**	1.16
11	It takes much too long to control the BCI application	**4.14**	1.83	**4.20**	1.40
12	I would love to participate in other similar studies	**6.43**	0.76	**5.20**	1.62
13	I found it difficult to select the desired commands	**2.93**	1.90	**2.80**	1.23
14	I would gladly carry out more testing sessions with the BCI application	**6.00**	1.47	**4.80**	1.62
15	I did not find the flickering effect annoying	**4.07**	1.59	**5.10**	1.85
16	The duration of the evaluation session was too long	**2.14**	1.56	**3.60**	1.51
17	I would not need a manual for controlling Twitter and Telegram with this system	**4.93**	1.77	**5.90**	1.73
18	I am happy that the sessions are over	**4.07**	1.59	**4.90**	1.29
19	I think that this system could improve the social media accessibility	**5.86**	1.41	**6.40**	0.70
20	I became impatient during the sessions	**2.07**	1.69	**3.40**	1.51

Statements were ranked in a 7-point Likert scale, where 1 means a complete disagreement, 4 a neutral response, and 7 a complete agreement.

first glance. The first task was easily completed by all the participants (CS: 100% $\pm$ 0.0%; MDS: 88.1% $\pm$ 16.6%). The second task was also completed by all the participants, even though they reached lower accuracies (CS:

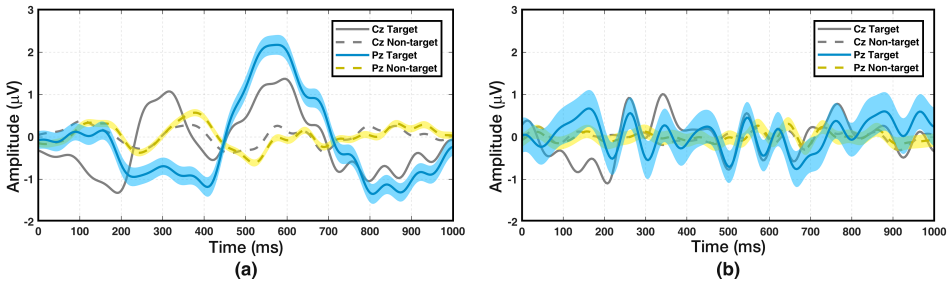


Figure 4: Event-related responses recorded in the first calibration session of two motor-disabled subjects: (a) M16, who could finish all tasks; and (b) M07, who was discarded due to its low classifier accuracy (<70%). Average curves of target stimuli (solid lines) and non-target stimuli (dashed lines) are depicted over the channel Pz (blue and yellow). Shaded areas indicate the 95% confidence interval of the aforementioned stimuli. Average curves over the channel Cz are also shown (grey). Note that a band-pass filter between 1–15 Hz has been applied for visualization purposes.

85.2% ± 20.7%; MDS: 89.0% ± 15.6%) and took three times more to finish than the first one. The third task was a struggle for M03, which could not finish it, probably because it was the first task that involved the use of both RCP matrices (CS: 90.0% ± 15.1%; MDS: 84.8% ± 12.5%). Like the previous one, the fourth task only was a problem for the same user, even though the duration increased appreciably (CS: 93.7% ± 7.5%; MDS: 77.5% ± 18.4%). The fifth task began to be challenging, and three MDS were not able to complete it (CS: 96.4% ± 4.6%; MDS: 77.2% ± 22.4%). Finally, the sixth task was by far the most difficult one, causing that five MDS could not finish it (CS: 89.0% ± 11.5%; MDS: 89.8% ± 14.9%). Note that, despite of the highest presumed difficulty of the latter, MDS accuracies in the sixth task are higher than that obtained in the fifth one. This is because the metrics are only computed for the users that could finish the task, reducing the performance variability, as indicated by the standard deviation. As revealed above, although all CS were able to finish all tasks, there were several MDS who faced problems to finish them. In particular, the two most challenging tasks involved the use of both matrices and spelling long sentences in order to communicate via Telegram chats. It was observed that a selection error often causes more mistakes thereafter, probably due to despondency. This issue could be solved by integrating a spelling dictionary or processing error-related potentials (ErrP) (Schalk et al., 2000).

Concerning the qualitative analysis, questionnaire results show that participants were quite satisfied with the BCI application. All the positive statements were valued above the neutral response (i.e., 4), and all the negative statements but two were valued below it. These statements were the 11th, which concerns the required time to control the application; and the 18th, which means that some users were slightly happy that the assessment sessions were over. The former discloses a request to increase the speed of the system. Nevertheless, the speed is directly related to the classifier performance, which depends on the user's calibration sessions. A more robust classifier, either because it would be based on a more sophisticated processing framework or because it would be trained with more data, could reach higher accuracies with fewer number of sequences, providing a faster navigation (Zhang et al., 2016). The latter reveals that the participation of several users implied an effort, a fact that should be taken into consideration when designing the tasks, their duration and the structure of the assessment sessions. However, users reported that they were willing to carry more sessions and to participate in further similar studies. Moreover, results show that these users did not experienced impatience, boredom, fatigue or stress. In addition, it is worthy to mention that the 6th statement was also valued below the neutral response for CS. This fact reveals that CS cannot imagine themselves using the BCI application in their daily life, which was expected because of their full physical and cognitive capabilities. Conversely, MDS do imagine themselves using the developed application as a daily tool, which reinforces the practicality of the system.

As pointed earlier, notwithstanding the growing popularity of smartphones, there are very few studies that have attempted to control their functionalities by integrating a BCI system. Table 5 shows these studies, which have been focused to dial numbers (Wang et al., 2011; Chi et al., 2012), accept incoming calls (Katona et al., 2014), call contacts (Campbell et al., 2010; Wang et al., 2011), play simple games (Wu et al., 2014), spell words (Obeidat et al., 2017; Elsawy et al., 2017) or open pre-installed apps and visualize the gallery (Elsawy and Eldawlatly, 2015). It is noteworthy that none of them has been focused on providing a high-level control of a smartphone, nor controlling social network functionalities. Moreover, the table 5 exposes one of the main drawbacks of the BCI literature,

Table 5: Comparison among state-of-the-art studies

Study	Control signal	EEG cap	Target SO	Processing	Main functionalities	*N*	Sub.	Accuracy[(1)]
Campbell et al. (2010)	P300	EPOC (Emotiv)	iOS	Mobile	Call contacts	3	CS	88.89%
Wang et al. (2011)	SSVEP	Custom headband	Cell phone	Computer	Dial numbers	10	CS	95.90%
Chi et al. (2012)	SSVEP	Custom dry electrode	Cell phone	Cell phone	Dial numbers	2	CS	89.00%
Katona et al. (2014)	Conc.	Mindset (Neurosky)	Windows phone	Headset	Accept/reject incoming calls	5	CS	75.00%
Wu et al. (2014)	Conc.	Mindset (Neurosky)	Android	Headset	Play a simple racing game	5	CS	-
Elsawy and Eldawlatly (2015)	P300	EPOC (Emotiv)	Android	Mobile	Open pre-installed apps and visualize the gallery	6 6	CS CS	79.17%[(2)] 87.5%
Elsawy et al. (2017)	P300	EPOC (Emotiv)	Android	Mobile	Spell words	6	CS	64.17%
Obeidat et al. (2017)	P300	EPOC (Emotiv)	Android	Mobile	Spell words	14	CS	90.00%
Present study	**P300**	**g.USBamp (g.Tec)**	**Android**	**Computer**	**Full asynchronous control of Twitter and Telegram**	**10 18**	**CS MDS**	**92.30% 80.60%**

"P300" refers to the P300 evoked potentials, "SSVEP" stands for steady-state visual evoked potentials, and "Conc." denotes a Neurosky concentration signal; "*N*" indicates the number of subjects; "CS" stands for control subjects, and "MDS" stands for motor-disabled subjects.
(1) Whether the study provides several accuracies for different experiments, the table shows the highest online reached performance. If accuracy is not provided directly, it is estimated from other data.
(2) The first accuracy belongs to the opening pre-intalled apps functionality, whereas the second one belongs to the visualizing application.

whose studies usually fail to prove the usability of their systems with end users. In fact, none of the aforementioned applications has been tested with motor-disabled users, who are the ones that would presumably benefit from them. It is also worthy to mention that none of these studies provides an asynchronous control, which implies that, in a real situation, an external supervisor should be present to pause the application when required. For this reason, one of the main objectives of this study is to evaluate our proposal with a population of 18 MDS in order to assess its usefulness to meet their daily communication needs.

Among the studies depicted in Table 5, P300 evoked potentials are the most prevalent control signals (Campbell et al., 2010; Elsawy and Eldawlatly, 2015; Obeidat et al., 2017; Elsawy et al., 2017). However, the customized Neurosky concentration metric is also used as an endogenous control signal (Katona et al., 2014; Wu et al., 2014), and steady-state visual evoked potentials (SSVEP) as exogenous ones (Wang et al., 2011; Chi et al., 2012). Even though the signal processing of the former is simple and can be handled by the headset itself, the Neurosky concentration signal can only be used to make dichotomous decisions. In other words, the systems of Katona et al. (2014) and Wu et al. (2014) could only discriminate two different EEG states, hindering the use of this signal for providing a high-level control of a complex system, such as the smartphones. Regarding the latter, it is worthy to mention that the SSVEP-based studies were both focused to dial numbers in cell phones (Wang et al., 2011; Chi et al., 2012). SSVEP signals are based on a mimetic mechanism: when the retina is excited by a visual stimulus that flickers at a constant frequency, the brain generates an oscillatory response at the same frequency (Wolpaw et al., 2002; Pastor et al., 2003; Capilla et al., 2011; Luck, 2014). The main advantage of the SSVEP signal is its exogenous nature, which makes a training phase unnecessary. Moreover, the signal also provides high performances, as the results show (Wang et al., 2011; Chi et al., 2012). However, the most reliable flickering frequencies belongs to the low beta band (i.e., 13–19 Hz) (Volosyak et al., 2011), which maximize the risk of epileptic seizures and visual fatigue (Pastor et al., 2003). Furthermore, the standardization of vertical refresh rate of LCD screens also restricts the number of simultaneously displayed frequencies (Volosyak et al., 2009). Therefore, the number of possible commands is limited. With regard to the P300-based studies, the use of a wireless headset with saline electrodes allows them to integrate a simple signal processing stage in the final devices (i.e., iOS or Android). However, although this solution favors the users' comfort and the practicality of the system, it also sets up a trade-off between portability and performance. In fact, the CS average accuracy of our study (92.30%) is higher than the ones reported in all these previous approaches, probably due to the use of: (i) gel-based active electrodes, (ii) a more complex signal processing module, and (iii) a larger stimulation screen. Significant differences have been found between our study outcomes and the results of the opening apps system of Elsawy and Eldawlatly (2015) (Wilcoxon Signed-rank Test, p-value $= 0.0088$); and the mobile speller of Elsawy et al. (2017) (Wilcoxon Signed-rank Test, p-value $= 0.0007$). The remaining P300-based studies do not provide unfolded accuracy results for each user and thus, statistical analysis could not be performed. Furthermore, it is worthy to mention that no comparison

with disabled subjects could have been made because of their lack of assessment with end users.

From the experimental outcomes, several insightful implications can be derived. On the one hand, this study may be considered as one of the first precursors of smartphone-based BCIs. As aforementioned, there are very few studies that have attempted to control mobile devices with BCI systems, and none of them was focused on providing a high-level control of a certain application. Our system provides a comprehensive control of two different social networks, covering all their functionalities and simultaneously reaching high accuracy results. To this end, users can select 72 different commands, arranged in two different RCP matrices. On the other hand, the present study has been tested with a population of both motor-disabled and control subjects and thus, the viability of the system has been demonstrated. Unfortunately, BCI-based studies usually fail to test their systems with real users, making it impossible to infer their reliability in a real context. Therefore, to the best of our knowledge, the present study is the first approach that has been proved its practicality to control a mobile BCI system by real users. These outcomes suggest that the developed system would be extended, in the near future, to assist individuals, companies or institutions that could be benefited from it. Consequently, personal autonomy and social integration of motor-disabled users could be improved, making an impact in their quality of life. To sum up, the main strengths of our proposal are:

i) Comprehensive control of Twitter and Telegram in Android platforms using brain signals.
ii) Ability to discriminate among a total of 72 different commands, arranged in two RCP matrices.
iii) Asynchronous control management by means of attention monitoring.
iv) Suitable performance accuracies.
v) Robustness, due to the evaluation with both control and motor-disabled populations.

Despite the results show that our BCI application allow users to successfully control Twitter and Telegram in an Android device, we can point out the following weaknesses:

i) Signal processing stage requires a laptop to be executed, which favors the reliability of the system, but impairs portability. Further research can overcome this limitation by using a wireless headset and integrating the processing stage into the final device.
ii) Asynchronous management is based on a wrapper method that depends on the LDA classifier and consequently, on the training performance of each user. Future endeavors must be focused on developing new asynchrony filter methods, such as SSVEP-based approaches independent of inter-session effects (Aloise et al., 2011; Pinegger et al., 2015; Wang et al., 2016; Jiao et al., 2017).
iii) Lack of despondency bypassing, causing a mistake to occasionally result in more errors in the following selections. A future research line could be aimed to implement a spelling dictionary or processing ErrPs to avoid extra selection errors (Cruz et al., 2018).
iv) Heterogeneous motor-disabled population. Although the application was tested with 18 MDS, and all of them can be considered end users of BCI systems, a future homogenization could be suitable for characterizing the performance of the system within a certain disease.

6. Conclusion

An asynchronous P300-based BCI system to control social networking applications of smartphones or tablets has been designed, developed and tested with both healthy and motor-disabled users. The system monitors the EEG signal of the user, while a RCP matrix containing the application commands flashes its rows and columns in order to generate P300 evoked potentials on the user's scalp. The selected commands are sent in real-time to the final Android device via Bluetooth, which interprets them and provides visual feedback to the user. The system has been tested with 10 CS and 18 MDS. The assessment was composed of two calibration stages and one evaluation session, where the users had to complete 6 different tasks, sorted by difficulty. Both quantitative and qualitative metrics were obtained, reaching average accuracies of 92.3% for CS and 80.6% for MDS. To the best of our knowledge, this is the first BCI study aimed to control social networking applications in a comprehensive way. Significant differences have been found among our accuracy results and that reported in other related studies, which obtained lower performances. Therefore, our P300-based BCI socializing system proves to be a suitable solution for motor-disabled users, allowing them to meet their daily communication needs.

In spite of the positive results, future research work can be suggested. Future endeavors should be aimed to: (i) embed the signal processing stage in the final device, (ii) design an asynchronous management independent of the classifier, (iii) implement a dictionary that suggests common words to the users based on their previous selections, (iv) process ErrPs to identify prediction errors and avoid wrong selections in real-time, and (v) test the application with a homogenized disabled population in order to study the performance within a certain disease.

Acknowledgments

This study was partially funded by projects TEC2014-53196-R and DPI2017-84280-R of 'Ministerio of Economía y Competitividad' and FEDER, the project "Análisis y correlación entre el genoma completo y la actividad cerebral para la ayuda en el diagnóstico de la enfermedad de Alzheimer" (Inter-regional cooperation program VA Spain-Portugal POCTEP 2014-202)

of the European Commission and FEDER, and project VA037U16 of the 'Junta de Castilla y León' and FEDER. V. Martínez-Cagigal and J. Gomez-Pilar were in receipt of a PIF-UVa grant of the University of Valladolid.

References

Aloise, F., Schettini, F., Aricò, P., Leotta, F., Salinari, S., Mattia, D., Babiloni, F., Cincotti, F., 2011. P300-based brain–computer interface for environmental control: an asynchronous approach. J. Neural Eng. 8 (2), 25025.

Atkinson, J., Campos, D., 2016. Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers. Expert Systems with Applications 47, 35–41.
URL http://dx.doi.org/10.1016/j.eswa.2015.10.049

Blankertz, B., Lemm, S., Treder, M., Haufe, S., Müller, K. R., 2011. Single-trial analysis and classification of ERP components - A tutorial. NeuroImage 56 (2), 814–825.
URL http://dx.doi.org/10.1016/j.neuroimage.2010.06.048

Blondet, M. V. R., Badarinath, A., Khanna, C., Jin, Z., 2013. A wearable real-time BCI system based on mobile cloud computing. International IEEE/EMBS Conference on Neural Engineering, NER, 739–742.

Campbell, A., Choudhury, T., Hu, S., Lu, H., Mukerjee, M. K., Rabbi, M., Raizada, R. D. S., 2010. NeuroPhone: brain-mobile phone interface using a wireless EEG headset. Proceedings of the Second ACM SIGCOMM workshop on Networking, systems, and applications on mobile handhelds (MobiHeld'10), 3–8.

Capilla, A., Pazo-Alvarez, P., Darriba, A., Campo, P., Gross, J., 2011. Steady-state visual evoked potentials can be explained by temporal superposition of transient event-related responses. PLoS ONE 6 (1).

Chen, L.-L., Zhao, Y., Ye, P.-F., Zhang, J., Zou, J.-Z., 2017. Detecting driving stress in physiological signals based on multimodal feature analysis and kernel classifiers. Expert Systems with Applications 85, 279–291.
URL http://dx.doi.org/10.1016/j.eswa.2017.01.040

Chi, Y. M., Wang, Y. T., Wang, Y., Maier, C., Jung, T. P., Cauwenberghs, G., 2012. Dry and noncontact EEG sensors for mobile brain-computer interfaces. IEEE Trans. Neural Syst. Rehabil. Eng. 20 (2), 228–235.

Corralejo, R., Nicolás-Alonso, L. F., Álvarez, D., Hornero, R., 2014. A P300-based brain–computer interface aimed at operating electronic devices at home for severely disabled people. Med. Biol. Eng. Comput. 52 (10), 861–872.

Cruz, A., Pires, G., Nunes, U. J., 2018. Double ErrP Detection for Automatic Error Correction in an ERP-Based BCI Speller. IEEE Transactions on Neural Systems and Rehabilitation Engineering 26 (1), 26–36.

Da Silveira, T. L., Kozakevicius, A. J., Rodrigues, C. R., 2016. Automated drowsiness detection through wavelet packet analysis of a single EEG channel. Expert Systems with Applications 55, 559–565.
URL http://dx.doi.org/10.1016/j.eswa.2016.02.041

Elsawy, A. S., Eldawlatly, S., 2015. P300-based applications for interacting with smart mobile devices. In: 7th Annual International IEEE EMBS Conference on Neural Engineering. pp. 166–169.

Elsawy, A. S., Eldawlatly, S., Taher, M., Aly, G. M., 2017. MindEdit: a P300-based text editor for mobile devices. Comput. Biol. Med. 80 (August 2016), 97–106.

Farwell, L. A., Donchin, E., 1988. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. Electroencephalogr. Clin. Neurophysiol. 70 (6), 510–523.

Hill, N. J., Wolpaw, J. R., 2016. Brain–Computer Interface. In: Reference Module in Biomedical Sciences. Elsevier.
URL http://www.sciencedirect.com/science/article/pii/B978012801238399322X

International Data Corporation, 2017. IDC quarterly mobile phone tracker. Tech. rep.
URL https://www.idc.com/tracker

Ipsos MORI, Google, 2017. Something for everyone: why the growth of mobile apps is good news for brands. Tech. rep.

Jasper, H. H., 1958. The ten twenty electrode system of the international federation. Electroencephalography and clinical neurophysiology 10, 371–375.

Jayabhavani, G. N., Raajan, N. R., Rubini, R., 2013. Brain mobile interfacing (BMI) system embedded with wheelchair. 2013 IEEE Conference on Information and Communication Technologies, ICT 2013 (Ict), 1129–1133.

Jiao, Y., Zhang, Y., Wang, Y., 2017. A Novel Multilayer Correlation Maximization Model for Improving CCA-Based Frequency Recognition in SSVEP Brain – Computer Interface. International Journal of Neural Systems 27 (8), 1–14.

Katona, J., Peter, D., Ujbanyi, T., Kovari, A., 2014. Control of incoming calls by a Windows Phone based Brain Computer Interface. 15th IEEE International Symposium on Computational Intelligence and Informatics (CINTI 2014), 121–125.

Keinosuke, F., 1990. Introduction to statistical pattern recognition. Academic Press Inc.

Kemp, S., 2017. Digital in 2017: global overview. Tech. rep.

Kleih, S. C., Nijboer, F., Halder, S., Kübler, A., 2010. Motivation modulates the P300 amplitude during brain-computer interface use. Clin. Neurophysiol. 121 (7), 1023–1031.

Krusienski, D. J., Sellers, E. W., Cabestaing, F., Bayoudh, S., McFarland, D. J., Vaughan, T. M., Wolpaw, J. R., 2006. A comparison of classification techniques for the P300 Speller. J. Neural Eng. 3 (4), 299–305.

Krusienski, D. J., Sellers, E. W., McFarland, D. J., Vaughan, T. M., Wolpaw, J. R., 2008. Toward enhanced P300 speller performance. J. Neurosci. Methods 167 (1), 15–21.

Kübler, A., Birbaumer, N., 2008. Brain-computer interfaces and communication in paralysis: Extinction of goal directed thinking in completely paralysed patients? Clin. Neurophysiol. 119 (11), 2658–2666.

Kübler, A., Kotchoubey, B., Kaiser, J., Birbaumer, N., Wolpaw, J. R., 2001. Brain-computer communication: Unlocking the locked in. Psychol. Bull. 127 (3), 358–375.

Kübler, A., Nijboer, F., Birbaumer, N., 2007. Brain-Computer Interfaces for communication and motor control – perspectives on clinical application. In: Toward Brain-Computer Interfacing, 1st Edition. MA: The MIT Press, pp. 373–391.

Likert, R., 1932. A technique for the measurement of attitudes. Arch. Psychol.

Luck, S. J., 2014. An introduction to the event-related potential technique. MIT press.

Ma, J., Zhang, Y., Cichocki, A., Matsuno, F., 2015. A novel EOG/EEG hybrid human-machine interface adopting eye movements and ERPs: Application to robot control. IEEE Transactions on Biomedical Engineering 62 (3), 876–889.

Martínez-Cagigal, V., Gomez-Pilar, J., Álvarez, D., Hornero, R., 2017a. An asynchronous P300-based Brain-Computer Interface web browser for severely disabled people. IEEE Trans. Neural Syst. Rehabil. Eng. 25 (8), 1332–1342.

Martínez-Cagigal, V., Hornero, R., 2017. P300-based Brain-Computer Interface channel selection using swarm intelligence. Rev. Iberoam. Autom. In. 14 (4), 372–383.

Martínez-Cagigal, V., Núñez, P., Hornero, R., 2017b. Spectral Regression Kernel Discriminant Analysis for P300 Speller Based Brain-Computer Interfaces. In: Converging Clinical and Engineering Research on Neurorehabilitation II. Biosystems & Biorobotics. Springer, Segovia, Spain, pp. 789–793.

Narsky, I., Porter, F. C., 2013. Statistical analysis techniques in particle physics: fits, density estimation and supervised learning. John Wiley & Sons.

Obeidat, Q., Campbell, T., Kong, J., 2017. Spelling with a small mobile Brain-Computer Interface in a moving wheelchair. IEEE Trans. Neural Syst. Rehabil. Eng. 4320 (c), 1.

Pastor, M. A., Artieda, J., Arbizu, J., Valencia, M., Masdeu, J. C., 2003. Human cerebral activation during steady-state visual-evoked responses. J. Neurosci. 23 (37), 11621–11627.

Picton, T. W., 1992. The P300 wave of the human event-related potential. J. Clin. Neurophysiol. 9 (4), 456–479.

Pinegger, A., Faller, J., Halder, S., Wriessnegger, S. C., Müller-Putz, G. R., 2015. Control or non-control state: that is the question! An asynchronous visual P300-based BCI approach. J. Neural Eng. 12 (1),

14001.
Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N., Wolpaw, J. R., 2004. BCI2000: A general-purpose brain-computer interface (BCI) system. IEEE Trans Biomed. Eng. 51 (6), 1034–1043.
Schalk, G., Wolpaw, J. R., McFarland, D. J., Pfurtscheller, G., 2000. EEG-based communication: presence of an error potential. Clin. Neurophysiol. 111 (12), 2138–2144.
Sellers, E. W., Donchin, E., 2006. A P300-based brain-computer interface: initial tests by ALS patients. Clin. Neurophysiol. 117 (3), 538–548.
Speier, W., Arnold, C., Pouratian, N., 2013. Evaluating true BCI communication rate through mutual information and language models. PLoS ONE 8 (10).
Townsend, G., LaPallo, B. K., Boulay, C. B., Krusienski, D. J., Frye, G. E., Hauser, C. K., Schwartz, N. E., Vaughan, T. M., Wolpaw, J. R., Sellers, E. W., 2010. A novel P300-based brain-computer interface stimulus presentation paradigm: moving beyond rows and columns. Clin. Neurophysiol. 121 (7), 1109–1120.
Volosyak, I., Cecotti, H., Gräser, A., 2009. Optimal visual stimuli on LCD screens for SSVEP based brain-computer interfaces. 2009 4th International IEEE/EMBS Conference on Neural Engineering, NER '09, 447–450.
Volosyak, I., Valbuena, D., Lüth, T., Malechka, T., Gräser, A., 2011. BCI demographics II: how many (and what kinds of) people can use a high-frequency SSVEP BCI? IEEE Trans. Neural Syst. Rehabil. Eng. 19 (3), 232–239.
Wang, H., Zhang, Y., Waytowich, N. R., Krusienski, D. J., Zhou, G., Jin, J., Wang, X., Cichocki, A., 2016. Discriminative Feature Extraction via Multivariate Linear Regression for SSVEP-Based BCI. IEEE Transactions on Neural Systems and Rehabilitation Engineering 24 (5), 532–541.
Wang, Y.-T., Wang, Y., Jung, T.-P., 2011. A cell-phone-based brain-computer interface for communication in daily life. J. Neural Eng. 8, 25018.
Wolpaw, J. R., Birbaumer, N., Heetderks, W. J., McFarland, D. J., Peckham, P. H., Schalk, G., Donchin, E., Quatrano, L. A., Robinson, C. J., Vaughan, T. M., 2000. Brain-computer interface technology: a review of the first international meeting. IEEE Trans. Rehabil. Eng. 8 (2), 164–173.
Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., Vaughan, T. M., 2002. Brain-computer interfaces for communication and control. Clinical Neurophysiology 113 (6), 767–91.
World Health Organization, 2011. World report on disability. Tech. rep., World Health Organization, Switzerland.
Wu, G., Xie, Z., Wang, X., 2014. Development of a mind-controlled Android racing game using a brain computer interface (BCI). 2014 4th IEEE International Conference on Information Science and Technology, 652–655.
Yuan, P., Gao, X., Allison, B., Wang, Y., Bin, G., Gao, S., 2013. A study of the existing problems of estimating the information transfer rate in online brain-computer interfaces. Journal of Neural Engineering 10 (2).
Zammouri, A., Ait Moussa, A., Mebrouk, Y., 2018. Brain-computer interface for workload estimation: Assessment of mental efforts in learning processes. Expert Systems with Applications 112, 138–147. URL https://doi.org/10.1016/j.eswa.2018.06.027
Zhang, Y., Zhou, G., Jin, J., Zhao, Q., Wang, X., Cichocki, A., 2016. Sparse Bayesian Classification of EEG for Brain-Computer Interface. IEEE Transactions on Neural Networks and Learning Systems 27 (11), 2256–2267.

Appendix B

Scientific achievements

B.1 Publications

B.1.1 Papers indexed in the JCR

1. **Víctor Martínez-Cagigal**, Javier Gomez-Pilar, Daniel Álvarez, Roberto Hornero, "An Asynchronous P300-Based Brain-Computer Interface Web Browser for Severely Disabled People", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25 (8), pp. 1332–1342, September, 2017, DOI: 10.1109/TNSRE.2016.2623381.

2. **Víctor Martínez-Cagigal**, Roberto Hornero, "P300-Based Brain-Computer Interface Channel Selection using Swarm Intelligence", *Revista Iberoamericana de Automática e Informática Industrial*, vol. 14 (4), pp. 372–383, October, 2017, DOI: 10.1016/J.RIAI.2017.07.003.

3. **Víctor Martínez-Cagigal**, Eduardo Santamaría-Vázquez, Roberto Hornero, "Asynchronous Control of P300-based Brain–Computer Interfaces using Sample Entropy", *Entropy*, vol. 21 (3), pp. 230, February, 2019, DOI: 10.3390/E21030230.

4. **Víctor Martínez-Cagigal**, Eduardo Santamaría-Vázquez, Javier Gomez-Pilar, Roberto Hornero, "Towards an Accessible Use of Smartphone-Based Social Networks through Brain-Computer Interfaces", *Expert Systems With Applications*, vol. 120, pp. 155–166, April, 2019, DOI: 10.1016/J.ESWA.2018.11.026.

5. **Víctor Martínez-Cagigal**, Eduardo Santamaría-Vázquez, Roberto Hornero, "Brain–Computer Interface Channel Selection Optimization using Meta-heuristics and Evolutionary Algorithms", *Applied Soft Computing*, (Under Review), 2020.

6. Eduardo Santamaría-Vázquez, **Víctor Martínez-Cagigal**, Javier Gomez-Pilar, Roberto Hornero, "Asynchronous Control of ERP-based BCI Spellers Using Steady-State Visual Evoked Potentials Elicited by Peripheral Stimuli", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27 (9), pp. 1883–1892, August, 2019, DOI: 10.1109/TNSRE.2019.2934645.

7. Eduardo Santamaría-Vázquez, **Víctor Martínez-Cagigal**, Fernando Vaquerizo-Villar, Roberto Hornero, "EEG-Inception: A Novel Deep Convolutional Neural Network for Assistive ERP-based Brain-Computer Interfaces", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, (Under Review), 2020.

B.1.2 Book chapters

1. **Víctor Martínez-Cagigal**, Eduardo Santamaría-Vázquez, Javier Gomez-Pilar, Roberto Hornero, "A brain–computer interface web browser for multiple sclerosis patients", in *Neurological Disorders and Imaging Physics, Volume 2: Engineering and Clinical Perspectives of Multiple Sclerosis*, ISBN-10: 0750317604, pp. 12:1–12:31, Bristol, United Kingdom: Institute of Physics Publishing, Editors: Ayman El-Baz and Jasjit S. Suri, 2019, DOI: 10.1088/978-0-7503-1762-7CH12.

B.1.3 International conferences

1. **Víctor Martínez-Cagigal**, Rebeca Corralejo, Javier Gomez-Pilar, Daniel Álvarez, Roberto Hornero, "Diseño, Desarrollo y Evaluación de un Navegador Web basado en Potenciales P300 mediante Brain-Computer Interface Orientado a Personas con Grave Discapacidad", *VI Congreso Internacional de Diseño, Redes de Investigación y Tecnología para Todos (DRT4ALL 2015)*, ISBN: 978-84-7993-277-0, pp. 343–368, Madrid (Spain), September 23 - September 25, 2015.

2. **Víctor Martínez-Cagigal**, Pablo Núñez, Roberto Hornero, "Spectral Regression Kernel Discriminant Analysis for P300 Speller Based Brain-Computer Interfaces", *International Conference on Neurorehabilitation*

(ICNR 2016), ISBN: 978-3-319-46668-2, pp. 789–793, La Granja, Segovia (Spain), October 18 - October 21, 2016, DOI: 10.1007/978-3-319-46669-9_-129.

3. **Víctor Martínez-Cagigal**, Roberto Hornero, "A Binary Bees Algorithm for P300-Based Brain-Computer Interfaces Channel Selection", *14th International Work-Conference on Artificial Neural Networks (IWANN 2017)*, ISBN: 978-3-319-59146-9, pp. 453–463, Cádiz (Spain), June 14 - June 16, 2017, DOI: 10.1007/978-3-319-59147-6_39.

4. Javier Gomez-Pilar, **Víctor Martínez-Cagigal**, Roberto Hornero, "Neurocognitive Training by means of a Motor Imagery-Based Brain Computer Interface in the Elderly", *6th International Conference on Cognitive Neurodynamics (ICCN 2017)*, pp. 46, Carmona (Spain), August 1 - August 5, 2017.

5. **Víctor Martínez-Cagigal**, Eduardo Santamaría-Vázquez, Roberto Hornero, "A Novel Hybrid Swarm Algorithm for P300-Based BCI Channel Selection", *World Congress on Medical Physics & Biomedical Engineering (IUPESM 2018)*, ISBN: 978-981-10-9022-6, pp. 41–45, Prague (Czech Republic), June 3 - June 8, 2018, DOI: 10.1007/978-981-10-9023-3_8.

6. **Víctor Martínez-Cagigal**, Eduardo Santamaría-Vázquez, Roberto Hornero, "Controlling a Smartphone with Brain-Computer Interfaces: A Preliminary Study", *X Conference on Articulated Motion and Deformable Objects (AMDO 2018)*, ISBN: 978-3-319-94543-9, pp. 34–43, Palma de Mallorca (Spain), July 12 - July 13, 2018, DOI: 10.1007/978-3-319-94544-6_4.

7. Eduardo Santamaría-Vázquez, **Víctor Martínez-Cagigal**, Javier Gomez-Pilar, Roberto Hornero, "Deep learning architecture based on the combination of convolutional and recurrent layers for ERP-based brain-computer interfaces", *XV Mediterranean Conference on Medical and Biological Engineering and Computing (MEDICON 2019)*, ISBN: 978-3-030-31635-8, pp. 1844–1852, Coimbra (Portugal), September 26 - September 28, 2019, DOI: 10.1007/978-3-030-31635-8_224.

8. **Víctor Martínez-Cagigal**, Reinmar J. Kobler, Valeria Mondini, Roberto Hornero, Gernot R. Müller-Putz, "Non-linear online low-frequency EEG decoding of arm movements during a pursuit tracking task", *42nd Annual International Conference of the IEEE Engineering in Medicine and*

Biology Society (EMBC'2020), ISBN: 978-1-7281-1990-8, pp. 2981–2985, Virtual conference (Canada), July 20 - July 24, 2020, DOI: 10.1109/EMBC44109.2020.9175723.

9. **Víctor Martínez-Cagigal**, Eduardo Santamaría-Vázquez, Roberto Hornero, "A Portable P300-based Brain–Computer Interface as an Alternative Communication Device", *5th International Conference on Neurorehabilitation (ICNR2020)*, Virtual conference (Spain), October 13 - October 16, 2020.

10. Eduardo Santamaría-Vázquez, **Víctor Martínez-Cagigal**, Daniel Rodríguez, Jaime Finat, Roberto Hornero, "Preventing Cognitive Decline in Elderly Population through Neurofeedback Training: A Pilot Study", *5th International Conference on Neurorehabilitation (ICNR2020)*, Virtual conference (Spain), October 13 - October 16, 2020.

B.1.4 National conferences

1. Javier Gomez-Pilar, Rebeca Corralejo,**Víctor Martínez-Cagigal**, Daniel Álvarez, Roberto Hornero, "Análisis de los cambios espectrales del EEG producidos por el entrenamiento neurocognitivo mediante una interfaz cerebro-ordenador",*7º Simposio CEA de Bioingeniería (CEA 2015)*, ISSN: 2341-4243, pp. 15–21, Málaga (Spain), June 25 - June 26, 2015.

2. **Víctor Martínez-Cagigal**, Javier Gomez-Pilar, Daniel Álvarez, Roberto Hornero, "Navegador Web BCI asíncrono controlado mediante potenciales P300", *8º Simposio CEA de Bioingeniería (CEA 2016)*, ISSN: 2341-4243, pp. 1–8, Madrid (Spain), June 23 - June 24, 2016.

3. **Víctor Martínez-Cagigal**, Javier Gomez-Pilar, Daniel Álvarez, Eduardo Santamaría-Vázquez, Roberto Hornero, "Sistema Brain-Computer Interface de Navegación Web Orientado a Personas con Grave Discapacidad", *XXXVIII Jornadas de Automática (JA 2017)*, ISBN: 978-84-16664-74-0, pp. 313–319, Gijón (Spain), September 6 - September 8, 2017.

4. **Víctor Martínez-Cagigal**, Roberto Hornero, "Multi-Objective Optimization for P300-Based Channel Selection", *9º Simposio CEA de Bioingeniería (CEA 2017)*, ISSN: 2341-4243, pp. 73–78, Barcelona (Spain), June 6 - June 7, 2017.

5. **Víctor Martínez-Cagigal**, Eduardo Santamaría-Vázquez, Roberto Hornero, "Interfaz Cerebro–Ordenador para el Control de las Funcionalidades de un Teléfono Móvil", *XXXVI Congreso Anual de La Sociedad Espñola de Ingeniería Biomédica (CASEIB 2018)*, ISBN: 978-84-09-06253-9, pp. 61–64, Ciudad Real (Spain), November 21 - November 23, 2018.

6. Eduardo Santamaría-Vázquez, **Víctor Martínez-Cagigal**, Roberto Hornero, "MEDUSA: Una Nueva Herramienta Para El Desarrollo De Sistemas Brain-Computer Interface Basada en Python", *10º Simposio CEA de Bioingeniería (CEA 2018)*, ISSN: 2341-4243, pp. 97–102, Madrid (Spain), July 2 - July 3, 2018.

7. Eduardo Santamaría-Vázquez, **Víctor Martínez-Cagigal**, Javier Gomez-Pilar, Roberto Hornero, "Control asíncrono de sistemas BCI basados en ERP mediante la detección de potenciales evocados visuales de estado estable provocados por los estímulos periféricos del paradigma oddball", *11º Simposio CEA de Bioingeniería (CEA 2019)*, pp. 1–11, Valencia (Spain), July 18 - July 19, 2019.

8. **Víctor Martínez-Cagigal**, Reinmar J. Kobler, Valeria Mondini, Roberto Hornero, Gernot R. Müller-Putz, "Decodificación no lineal de los movimientos de la mano en tiempo real mediante un sistema Brain–Computer Interface", *XXXVIII Congreso Anual de La Sociedad Española de Ingeniería Biomédica (CASEIB 2020)*, Virtual conference (Spain), November 25 - November 27, 2020.

B.2 International internship

Three-month research internship at the Institute of Neural Engineering (INE), Graz University of Technology (TU Graz), Graz, Austria.

i. **Purpose of the internship**
   The main purpose of the research stay was to investigate whether it was possible to decode continuous movements of the upper-limb from the EEG, aiming at controlling an assistive robotic arm in real-time. The work was part of the project "Feel Your Reach" of the Europe Research Council (ERC), focused on developing EEG-based BCI-controlled neuroprosthesis. To carry out that objective, the developed study encompassed: (1) a state-of-the-art revision of non-linear decoding algorithms, (2) coding of different Kalman

filters in MATLAB®; (3) integration of the final algorithms into an online paradigm, (4) pilot testing of the system with 2 HS, (5) real testing with 5 HS, (6) analysis of the results (quantitative, qualitative, brain sources), and (7) presentation.

ii. Methodological summary

In a nutshell, the experimental paradigm that was carried out was the following. The subject was comfortably seated in front of a screen that displayed a white dot following a pre-defined trajectory in 2D. Three main devices comprised the BCI: (i) the EEG acquisition equipment (64 channels, 500 Hz); (ii) the LeapMotion (LM), which monitored the position of user's hand using infrared image processing; and (iii) the JACO robot, an assistive robotic arm. The latter was controlled by both the EEG and LM, and the task of the user was to follow the trajectory with the JACO's hand as best as they could. At the beginning, JACO was entirely LM-controlled. After a few trials of calibration, the BCI began to rely on the EEG, reducing the control of LM until it became fully EEG-controlled. Motion and ocular artifacts were corrected. Non-linear decoding was performed using partial least squares (PLS) regression and square-root unscented Kalman filtering (SQ-UKF). Correlations between actual and decoded movements were generally above chance level, suggesting that the proposed system was suitable for decoding real-time arm movements.

iii. Quality indicators of the institution

Founded in 1811 by Archduke John of Austria, the TU Graz is one of the main universities of Austria. This public university is part of the network *Austrian Universities of Technology* and reaches suitable positions in international rankings (Shanghai: 151-200, Times High Education: 351-400). The INE of TU Graz has become an institution of international renown in the field of BCI systems. In fact, the well-known 'Graz BCI' was the first BCI system to debut in the history of neuroscience, 20 years ago. Since then, the INE has performed a strong research focused on developing assistive BCI applications. The head of the INE, Dr. Gernot M"uller-Putz, who was also the supervisor of the internship, was the principal investigator of two European projects that took place during the research stay ("MoreGrasp" EU Project; "Feel Your Reach", ERC-COG H2020). In the last 10 years, the INE has participated in 10 international and 8 national projects; published more than 193 JCR articles, 419 conferences, 36 book chapters;

and defended 9 doctoral theses. In light of this intense research, the INE organizes the "Graz BCI Conference" once every three years since 2006, which attracts BCI researchers from around the world.

B.3 Awards and honors

03/2016: **Secondary Award in the University-Business Challenge (TCUE2015-16)**, for the project entitled "Navegador web accesible controlado mediante BCI basado en potenciales P300".

09/2016: **Prize for Special Achievement in the Master's Degree in ICT Research**, due to the obtaining of the highest marks of his class.

12/2017: **Award in Innovative Solutions for the Improvement of the Quality of Life** ('Ageing' group modality), for the project entitled "Plataforma Brain-Computer Interface de Entrenamiento Cognitivo para Atenuar los Efectos del Envejecimiento", conducted by Víctor Martínez-Cagigal, Javier Gomez-Pilar and Roberto Hornero.

09/2019: **Second prize in the IFMBE Scientific Challenge at the MEDICON 2019**, for the project entitled "Deep learning architecture based on the combination of convolutional and recurrent layers for ERP-based brain-computer interfaces", conducted by Eduardo Santamaría-Vázquez, Víctor Martínez-Cagigal, Javier Gomez-Pilar, and Roberto Hornero.

Apéndice C

Resumen en castellano

C.1 Introducción

La idea de establecer una conexión entre nuestros cerebros y el entorno, así como la posibilidad de controlar dispositivos mediante nuestras señales cerebrales, ha fascinado a la humanidad durante el último siglo. El descubrimiento del electroencefalograma (EEG) por Hans Berger, el trabajo inicial de Jacques Vidal y los progresos en la investigación neurocientífica actual poco a poco hacen esta idea cada vez más factible. Hoy en día, la ciencia a ficción empieza a convertirse en realidad. Durante los últimos 25 años, numerosos grupos de investigación han dedicado esfuerzos a decodificar señales neuronales y provocar el desarrollo de los sistemas *brain–computer interface* (BCI), entendidos como sistemas de comunicación que traducen las intenciones del usuario en comandos de un dispositivo externo. Invasivos y no invasivos, dependientes e independientes, exógenos y endógenos, activos y pasivos, síncronos y asíncronos; existen multitud de sistemas BCI que constantemente se ven mejorados por la academia y, sin embargo, la mayor parte de ellos aún no son lo suficientemente fiables como para abandonar los laboratorios y permitir su uso en entornos reales.

Desde un punto de vista práctico, un sistema BCI debería ser idealmente no invasivo, portátil, fiable, cómodo y robusto ante distintos entornos y artefactos externos. Por esta razón, las técnicas invasivas como la electrocorticografía (ECoG) o la tomografía por emisión de positrones (PET); y las voluminosas e inasequibles, como la magnetoencefalografía (MEG), la imagen por resonancia magnética funcional (fMRI) o la espectroscopía de infrarrojo cercano funcional (fNIRS) normalmente quedan relegadas al campo de la investigación. En esta Tesis Doctoral,

por tanto, se tratan los sistemas BCI basados en EEG, debido a su naturaleza no invasiva, su bajo coste y su portabilidad. En el EEG se colocan una serie de electrodos sobre el cuero cabelludo del usuario, permitiendo recoger la actividad eléctrica conjunta de millones de neuronas al mismo tiempo. No obstante, los electrodos solamente son sensibles a la actividad de neuronas superficiales organizadas perpendicularmente (i.e., giros), ignorando completamente aquellas que forman los surcos cerebrales, así como las presentes en estructuras subcorticales. Este hándicap hace del EEG una señal burda, donde las intenciones del usuario se encuentran sepultadas bajo ruido eléctrico y actividades neurológicas no relacionadas. Decodificar las intenciones de usuario directamente de la señal EEG, por tanto, es muy complejo técnicamente. Para ello, los sistemas BCI dependen de las señales de control: estrategias que provocan cambios en el EEG detectables a través de un procesado de la señal adecuado. Entre ellas destacan los ritmos sensoriomotores (SMR), los potenciales corticales lentos (SCP), los potenciales evocados visuales de estado estable (SSVEP) y los potenciales evocados P300.

Por una parte, los SMR y SCP se basan en autorregular la actividad cerebral de manera voluntaria. Los usuarios, por tanto, requieren de un entrenamiento exhaustivo para aprender a realizarlo, e incluso muchos de ellos no llegan a adquirir un control suficiente para trabajar con un sistema BCI. Este aspecto, sumado al hecho de que estas señales normalmente solo permiten tomar decisiones dicotómicas, restringen su uso práctico a la investigación o a terapias basadas en *neurofeedback*. Por otra parte, los SSVEP y los potenciales evocados P300 no requieren entrenamiento por parte de los usuarios, puesto que se basan en provocar respuestas naturales del cerebro ante distintos estímulos externos. Ambas permiten alcanzar precisiones superiores al 90 % con facilidad, siendo éstas las señales de control más adecuadas para sistemas BCI orientados a controlar aplicaciones o dispositivos externos. La generación de SSVEPs se basa en mostrar un conjunto de opciones que parpadeen a distintas frecuencias, provocando que la frecuencia de la opción a la que atienda el usuario se vea reflejada en el espectro de la señal EEG. No obstante, las frecuencias más fiables pertenecen a la banda beta baja (i.e., 13–19 Hz), lo que incrementa la fatiga visual y maximiza el riesgo de ataques de epilepsia fotosensible. Asimismo, el número de comandos posibles está ligeramente limitado por la frecuencia de muestreo de las pantallas LCD. Los potenciales evocados P300, por el contrario, se basan en la aparición de estímulos inesperados para los usuarios, maximizando su precisión cuantas más opciones posibles existan. Este aspecto ha provocado que los sistemas BCI basados en P300 estén siendo utilizados actualmente por personas con graves discapacidades motoras en su vida diaria. Dado

que la presente Tesis Doctoral busca contribuir en la literatura de los sistemas BCI desde un punto de vista práctico, todos los estudios incluidos en este compendio de publicaciones tratan sobre sistemas basados en potenciales evocados P300, debido a las razones mencionadas.

A pesar del gran avance de los sistemas BCI no invasivos en las últimas décadas, existen cuatro limitaciones principales que dificultan su uso práctico por parte de personas con grave discapacidad, relacionadas con: (1) el *hardware*, (2) la fiabilidad, (3) la validación, y (4) la sincronía. El *hardware* actual limita enormemente la portabilidad y comodidad de sistema. Idealmente, un sistema BCI debería ser totalmente portable, inalámbrico, ser fácil de configurar, no requerir gel conductor, trabajar varias horas sin mantenimiento y funcionar correctamente en los distintos escenarios de la vida real. La fiabilidad hace referencia a la alta variación entre sesiones y sujetos de la precisión de sistema. Aunque normalmente mejora con la práctica, el rendimiento nunca llega a ser similar al de un control muscular. Otro problema es la ausencia de validación. Muchos de los estudios de sistemas BCI en la literatura carecen de una evaluación con sujetos reales (i.e., personas con grave discapacidad) y, por tanto, su viabilidad no puede ser asegurada. Es conocido ampliamente que los sujetos de control obtienen mejores rendimientos que los sujetos con grave discapacidad, siendo inadecuado generalizar los resultados. Finalmente, cabe destacar que los sistemas BCI deben evolucionar hacia un estado asíncrono. Los sistemas BCI tradicionales son síncronos, es decir, no monitorizan la atención de usuario y provocan la generación de respuestas incluso cuando éste no desea generarlas. En la práctica este efecto se traduce en envíos involuntarios de comandos cuando el usuario no está prestando atención al sistema. La mayor parte de las aplicaciones de asistencia, por tanto, no proveen al usuario de un control total del sistema, requiriendo la presencia constante de un supervisor.

En esta Tesis Doctoral se presenta un compendio de cuatro publicaciones indexadas en el *Journal Citation Reports* (JCR) entre los años 2017 y 2020. Dos de ellas se centran en el procesado de señal: (1) asincronía mediante métricas basadas en entropía (Martínez-Cagigal et al., 2019b), y (2) metaheurísticas para seleccionar canales relevantes (Martínez-Cagigal et al., 2020); mientras que en el resto se detalla el diseño, desarrollo y evaluación de aplicaciones BCI asíncronas que permiten el control de: (3) un navegador web (Martínez-Cagigal et al., 2017), y (4) redes sociales (Twitter y Telegram) en un *smartphone* (Martínez-Cagigal et al., 2019a).

C.2 Hipótesis y objetivos

A pesar del creciente interés científico en los sistemas BCI durante las últimas décadas, las limitaciones mencionadas los han relegado a un uso en laboratorios con propósitos académicos. La hipótesis de partida de esta Tesis Doctoral baraja la posibilidad de que *las limitaciones de los sistemas BCI que dificultan su aplicación práctica en entornos reales puedan atenuarse*. En particular, la sincronía puede abordarse dotando al sistema BCI de un control dependiente únicamente del usuario. Un control asíncrono puede alcanzarse siempre y cuando el sistema sea capaz de determinar si el usuario quiere seleccionar un comando (i.e., estado de control) o no (i.e., estado de no-control). En este sentido, se hipotetiza que *los potenciales evocados P300 no se presentan en el estado de no-control, disminuyendo la probabilidad a posteriori del clasificador*. Asimismo, dado que puede considerarse que el estado de control es más exigente que el de no-control, también se hipotetiza que *determinadas medidas basadas en entropía pueden caracterizar la irregularidad de las señales EEG y proveer información que ayude a discriminar entre ambos estados*. Con respecto al *hardware*, la selección de canales relevantes permitiría reducir el coste del equipo y el consumo de energía, así como incrementar la comodidad de usuario. En este aspecto se hipotetiza que *las metaheurísticas basadas en algoritmos evolutivos son capaces de hallar los canales relevantes para cada usuario en sistemas P300-BCI*, puesto que su utilidad resolviendo problemas de optimización complejos ha sido extensamente demostrada. Paralelamente, se espera que *estos métodos ayuden a evitar el exceso de dimensionalidad y maximizar la precisión del sistema*, contribuyendo a la fiabilidad de mismo. Finalmente, cabe destacar que la viabilidad de los sistemas deberá ser evaluada mediante su aplicación con usuarios reales. Por esa razón, se hipotetiza si *un navegador web P300-BCI asíncrono es capaz de proveer a las personas con grave discapacidad de una tecnología viable para acceder a Internet*. Por extensión, también se hipotetiza si *un sistema P300-BCI asíncrono permite controlar plenamente distintas redes sociales (i.e., Twitter y Telegram) en un smartphone*. Para validar estas hipótesis, esta Tesis Doctoral propone *el uso de distintas metodologías para contribuir en el desarrollo de sistemas BCI prácticos en un entorno real*.

Definidas las hipótesis, el objetivo general es *diseñar, desarrollar y evaluar novedosas técnicas de procesado de señal y aplicaciones de asistencia para proveer sistemas P300-BCI a personas con grave discapacidad*. Para llevar a cabo este objetivo, se han alcanzado los siguientes objetivos específicos:

I. Revisar la bibliografía y los últimos avances relacionados con los sistemas

BCI no invasivos, poniendo un énfasis especial en métodos de selección de canales y gestión de la asincronía, así como en el desarrollo de aplicaciones de asistencia.

II. Construir una base de datos de señales EEG pertenecientes al estado de control y no-control para abordar el problema de la asincronía en sistemas P300-BCI, así como reclutar una población de usuarios con graves discapacidades motoras para validar las aplicaciones de asistencia.

III. Implementar los métodos más apropiados para optimizar los canales para cada usuario, discriminar entre los estados asíncronos e identificar potenciales P300; así como investigar la idoneidad de distintas mejoras.

IV. Diseñar y desarrollar dos aplicaciones asíncronas de asistencia para controlar: (1) un navegador web, y (2) redes sociales en un *smartphone*.

V. Evaluar la habilidad de los métodos seleccionados para optimizar los canales de cada usuario y para alcanzar un estado asíncrono en sistemas P300-BCI mediante su testeo en nuestra base de datos y otras de acceso público. Validar las aplicaciones de asistencia desarrolladas con la base de datos de sujetos con grave discapacidad, así como con sujetos de control.

VI. Aplicar análisis estadísticos de los resultados para evaluar la fiabilidad de cada una de las propuestas, así como comparar y discutir los resultados para extraer conclusiones oportunas, incluyendo una comparación exhaustiva con otros estudios relacionados.

VII. Diseminar los resultados principales y las conclusiones de los estudios en revistas JCR indexadas, así como en capítulos de libros y conferencias internacionales y nacionales.

C.3 Sujetos

Dado que los propósitos de los distintos artículos que forman el compendio varían, se han usado distintas bases de datos durante el proceso de realización de esta Tesis Doctoral. En la tabla C.1 se recogen las especificaciones de las bases de datos empleadas. Para el estudio de asincronía, se recogieron señales de 10 sujetos de control (HS) atendiendo e ignorando el *row-col paradigm* (RCP) (Martínez-Cagigal et al., 2019b). En el estudio de selección de canales se emplearon 3 bases de datos públicas con señales de equipos EEG de alta densidad. En cada una de

Cuadro C.1: Especificaciones de los sujetos de cada estudio.

Estudio	Pacientes	Sujetos de control	Paradigma	Nº de canales
Martínez-Cagigal et al. (2019b)	0	10	RCP	16
Martínez-Cagigal et al. (2020)	0	2	[1]RCP	64
	0	13	[2]CS	63
	0	12	[3]RSVP	61
Martínez-Cagigal et al. (2017)	16 MS	5	RCP	8
Martínez-Cagigal et al. (2019a)	18 MDS[4]	10	RCP	8

MS: esclerosis múltiple, MDS: sujetos con grave discapacidad motora, RCP: *row-col paradigm*, CS: *center speller*, RSVP: *rapid serial visual presentation*.
[1] 'BCI Competition III: dataset II' (Blankertz et al., 2006).
[2] 'Center Speller (008-2015)' (Treder et al., 2011).
[3] 'RSVP Speller (010-2015)' (Acqualagna and Blankertz, 2013).
[4] 1 accidente cerebrovascular, 2 lesiones de médula espinal, 5 ataxias de Friedreich, 5 parálisis cerebrales, 2 distrofias musculares.

ellas se emplearon paradigmas *oddball* distintos, con el fin de favorecer la generalización de los resultados: RCP (2HS), *center speller* (CS, 13HS) y *rapid serial visual presentation* (RSVP, 12HS) (Martínez-Cagigal et al., 2020). Las aplicaciones de asistencia no solo fueron evaluadas por sujetos de control, sino también por usuarios con graves discapacidades motoras, reclutados a través del Centro de Referencia Estatal de Discapacidad y Dependencia de San Andrés del Rabanedo (León). El sistema BCI para controlar el navegador web fue evaluado por 5HS y 16 usuarios con esclerosis múltiple (MS) (Martínez-Cagigal et al., 2017), mientras que la aplicación móvil de redes sociales se testeó con 10HS y 18 usuarios que presentaban distintas discapacidades motoras (MDS).

C.4 Métodos

La metodología de los distintos estudios comparte la misma estructura general, compuesta por las etapas de: (1) pre-procesado; (2) extracción, (3) selección y (4) clasificación de características; y (5) análisis estadístico. A esta estructura se le añaden los métodos de asincronía, selección de canales y la etapa de aplicación dependiendo del estudio en particular.

Como pre-procesado, todas las señales sufrieron un proceso de acondicionamiento para disminuir la presencia de artefactos mediante filtros frecuenciales (paso-banda 0.1–60 Hz, ranura a 50 Hz) y espaciales (referencia de media común) (Krusienski Dean J., 2012). La extracción de características se realizó a través de un proceso de decimación (f_d =20–25 Hz), seguido por un enventanado (0–800 ms,

z-scored baseline -200–0 ms) y una concatenación (Krusienski Dean J., 2012). Esta metodología busca extraer las características temporales más relevantes después de cada estímulo, con el fin de detectar posibles *event-related potentials* (ERP). La selección de características se realizó mediante una regresión paso-a-paso (SW) con criterios de inclusión y exclusión: $p < 0{,}10$ y $p > 0{,}15$, respectivamente (Jobson, 1991). Posteriormente, se empleó un análisis discriminante lineal (LDA) para determinar la probabilidad de aparición del P300 en cada observación (Bishop, 2006). Con el fin de determinar si los cambios en los resultados arrojan diferencias significativas, se aplicaron diversos test estadísticos no paramétricos. Para las comparaciones pareadas (i.e., dependientes), se empleó el test de rangos con signo de Wilcoxon; mientras que para las no pareadas (i.e., independientes) se utilizó el test U de Mann-Whitney (Narsky and Porter, 2013). Los p-valores de cada uno de los tests fueron adaptados contra el efecto de múltiples comparaciones mediante el método de Benjamini-Hochberg de corrección de la *false discovery rate* (FDR) (Benjamini and Hochberg, 1995).

Para alcanzar un estado asíncrono se emplearon dos técnicas distintas: (i) umbralización (Martínez-Cagigal et al., 2017, 2019a) y (ii) entropía muestral (Martínez-Cagigal et al., 2019b). En las aplicaciones de asistencia se desarrolló y evaluó el método de (i) umbralización. Esta técnica asume que los *scores* del clasificador LDA son más bajos cuando el usuario no presta atención a la estimulación visual (i.e., estado de no-control) que cuando sí lo hace (i.e., estado de control). Una vez registradas las señales de control y no-control, se introducen los *scores* de las celdas seleccionadas en una curva *receiver operating characteristic* (ROC) y se determina el valor de umbral óptimo que maximice el par sensibilidad-especificidad. Cada vez que el paradigma seleccione un comando, se compara su *score* con el umbral, enviándolo a la aplicación final si lo supera, o evitando la selección si no lo hace (Martínez-Cagigal et al., 2017). El método basado en (ii) entropía muestral, por el contrario, es totalmente independiente del clasificador. Dado que la estimación de la entropía es más precisa cuantas más muestras de la señal existan, la extracción de características realizada difiere de la estructura general. En este caso, la época i sería la señal decimada desde la primera estimulación hasta el último flash de la secuencia i. De esta manera, se asegura que para cada observación se consideran todas las muestras disponibles del comando actual. Posteriormente, se caracterizaron las señales de control y no-control mediante la aplicación de la entropía multiescala (MSE) (Costa et al., 2005). La optimización de los hiperparámetros se llevó a cabo a través de una validación cruzada dejando uno fuera. Tras la aplicación de la entropía muestral (SampEn), se utilizó un clasificador LDA

para monitorizar la atención del usuario (Martínez-Cagigal et al., 2019b).

La selección de canales se llevó a cabo a través de la aplicación de un conjunto de metaheurísticas mono y multiobjetivo. En todas ellas se buscó maximizar el rendimiento del sistema (i.e., área bajo la curva ROC) y simultáneamente minimizar el número de canales. Los algoritmos mono-objetivo empleados fueron: *genetic algorithm* (GA), *binary differential evolution* (BDE), y *binary particle swarm optimization* (BPSO); mientras que los multiobjetivo fueron: *non-sorting genetic algorithm II* (NSGA-II), *binary multi-objective PSO* (BMOPSO), y *strength pareto evolutionary algorithm 2* (SPEA2). No obstante, todas las métricas mencionadas tuvieron que ser adaptadas al contexto de la selección de canales en sistemas BCI, puesto que muchas de sus estrategias internas son subóptimas o incluso fútiles (e.g., distancias continuas o de multitud, funciones de transferencia, control del repositorio, etc.). Por esta razón, en Martínez-Cagigal et al. (2020) se diseñó el *dual-front sorting genetic algorithm* (DFGA), un algoritmo multiobjetivo específicamente diseñado para este problema de optimización que mezcla técnicas deterministas y estocásticas y evita todas las estrategias subóptimas para favorecer la convergencia.

C.5 Resultados y discusión

Los dos métodos de asincronía desarrollados han demostrado su utilidad tanto para discriminar entre estados de control y no-control en aplicaciones de asistencia (umbralización) como para caracterizar ambos estados en términos de complejidad y regularidad de las señales EEG (MSE, SampEn). La umbralización alcanzó precisiones de entrenamiento para los HS del 96.66 % y 96.74 % en la evaluación del navegador web y de la aplicación de redes sociales; mientras que para los usuarios con grave discapacidad se alcanzaron precisiones del 86.77 % y 84.31 %, respectivamente (Martínez-Cagigal et al., 2017, 2019a). Su implementación en aplicaciones de asistencia, por tanto, es viable. No obstante, el método depende totalmente del clasificador ERP, lo que conlleva registrar señales de ambos estados cada vez que se actualice el mismo. Este procedimiento, de hecho, se realiza con frecuencia debido a la gran variabilidad inter-sesión del EEG, lo que supone un incremento del tiempo de calibración. El método basado en la entropía muestral, por el contrario, es independiente del clasificador. Se ha observado una mayor irregularidad y complejidad en las señales de control, favoreciendo la discriminación entre los estados asíncronos, sobre todo en la zona prefrontal del córtex. Este método alcanzó una precisión del 94.4 %, con un coste computacional de 196.8 ms para 15 secuencias, lo

que lo convierte en una alternativa viable en tiempo real (Martínez-Cagigal et al., 2019b). Hasta donde llega nuestro conocimiento, no existen estudios previos que hayan tratado de caracterizar los estados asíncronos mediante métricas basadas en entropía. Otros estudios proponen técnicas de umbralización (Aloise et al., 2011; Breitwieser et al., 2016; Tang et al., 2018; Zhang et al., 2008), el uso de otras señales de control en paradigmas híbridos (Li et al., 2013; Panicker et al., 2010; Yu et al., 2017), o análisis espectrales (Ma and Qiu, 2018; Pinegger et al., 2015; Santamaría-Vázquez et al., 2019).

La optimización de canales relevantes mediante metaheurísticas basadas en algoritmos evolutivos han obtenido resultados con alta precisión para todas las bases de datos, superando de manera significativa al conjunto de 8 canales propuesto por Krusienski et al. (2008) (KRU) y al set de todos los electrodos (ALL). Los algoritmos mono-objetivo (GA, BDE, BPSO) alcanzaron, de media, un 91.47 % de precisión usando 13.38 canales (Martínez-Cagigal et al., 2020). Sin embargo, no ofrecen un control al supervisor sobre el número total de canales a seleccionar, a diferencia de los métodos multiobjetivo. Estos últimos retornan un conjunto de soluciones óptimas, cada cual con un número distinto de canales. Entre ellos, DFGA mostró una convergencia excelente, seguido por NSGA-II y SPEA2; y los tres métodos superaron a la solución KRU usando la mitad de canales (Martínez-Cagigal et al., 2020). BMOPSO, sin embargo, no convergió adecuadamente, cayendo en mínimos locales y mostrando una deficiencia en la optimización global de los objetivos. Se observó una gran variabilidad inter-sujeto en el conjunto de canales seleccionados, lo que pone de manifiesto la necesidad de optimizar individualmente el conjunto de canales para cada usuario, en lugar de usar la misma configuración para todos ellos. No obstante, los algoritmos mostraron una ligera tendencia a seleccionar canales sobre el lóbulo occipital, donde se encuentra el córtex visual primario (Martínez-Cagigal et al., 2020). La mayor parte de aproximaciones previas en la literatura emplean técnicas mono-objetivo, a pesar de la limitación mencionada (Arican and Polat, 2020; Gonzalez et al., 2013; Jin et al., 2010; Martínez-Cagigal and Hornero, 2017b; Perseh and Sharafat, 2012). Los estudios con técnicas multiobjetivo son escasos (Chaurasiya et al., 2017; Kee et al., 2015). No obstante, hasta donde llega nuestro conocimiento, ningún estudio ha testeado metaheurísticas en otros paradigmas distintos al RCP, ni tampoco han propuesto ningún algoritmo específicamente diseñado para la optimización de canales relevantes en sistemas BCI. DFGA, por tanto, constituye una aproximación novedosa y viable para aumentar el rendimiento, reducir el coste del equipo y favorecer la comodidad de los usuarios.

La primera de las aplicaciones de asistencia, el navegador web, resultó apropiada para su uso por parte de personas con grave discapacidad. La aplicación se controlaba mediante dos matrices RCP distintas: 'navegación', que permitía un control general de la aplicación mediante selecciones rápidas debido a su pequeño tamaño; y 'teclado', orientada a rellenar formularios y escribir texto (Martínez-Cagigal et al., 2017). La navegación se realizaba bajo una estrategia de etiquetado de nodos, asignando una codificación a cada uno de los elementos de la web. El usuario, por tanto, introducía la codificación numérica mediante las matrices RCP para seleccionar el nodo deseado. El testeo se llevó a cabo durante 2 sesiones de calibración (clasificador y umbral asíncrono) y 2 de evaluación (síncrona y asíncrona). Durante estas últimas se realizaron un total de 8 tareas con dificultad incremental, con el fin de evaluar la capacidad de los usuarios para controlar la aplicación. Se obtuvo una precisión media del 95.75 % para los HS y del 84.14 % para los sujetos MS, lo que demuestra la viabilidad de la aplicación (Martínez-Cagigal et al., 2017). Tres sujetos MS fueron descartados durante la calibración por sus bajas respuestas P300 (i.e., potenciales atenuados o nulos, latencias variables), probablemente debido a sus características clínicas. Este hecho pone de manifiesto la importancia de evaluar los sistemas BCI con sujetos reales, puesto que es bien sabido que los HS obtienen precisiones más altas que aquellas personas que presentan enfermedades neurodegenerativas. Aun así, las precisiones obtenidas por los MS superan significativamente las reportadas en estudios previos evaluados por sujetos con esclerosis lateral amiotrófica (Karim et al., 2006; Mugler et al., 2010). Asimismo, la precisión de los HS también supera otras aproximaciones previas (Mugler et al., 2010; Sirvent Blasco et al., 2012; Yu et al., 2012), sugiriendo que la inclusión de la etapa asíncrona favorece la aplicabilidad de los sistemas BCI en entornos asistenciales. Los cuestionarios reflejaron la satisfacción de los usuarios con el sistema, encontrándolo interesante, intuitivo y declarando que podrían imaginarse usándolo como herramienta durante su vida diaria.

La aplicación BCI móvil para controlar las redes sociales en el *smartphone* se desarrolló utilizando de las APIs de Twitter y Telegram. Al igual que para el navegador, se emplearon dos matrices RCP para controlar la aplicación. Se utilizó el etiquetado de nodos para todas aquellas funcionalidades que no podía ser controladas mediante comandos RCP. El testeo estuvo compuesto por dos sesiones de calibración (clasificador y umbral) y una de evaluación asíncrona, compuesta por 5 tareas que debían completar los usuarios. Un total de 4 participantes MDS se descartaron en la calibración por obtener precisiones inferiores al 70 %, causadas por respuestas P300 atenuadas (Martínez-Cagigal et al., 2019a). Se observó

una baja capacidad de atención sostenida, temblores involuntarios y nistagmo en algunos usuarios, lo cual afectó a la precisión del sistema. Este tipo de problemas suelen presentarse en las evaluaciones con usuarios reales, dificultando la generalización de los resultados de sistemas testeados únicamente con HS. Se obtuvieron precisiones medias del 92.3 % para los HS, y de 80.6 % para los MDS. Estos últimos también experimentaron una velocidad de selección de comandos significativamente inferior, puesto que el número de secuencias óptimo para ellos fue más alto. Los resultados de los cuestionarios reflejaron una satisfacción general con el sistema, indicando que no experimentaron fatiga, estés o aburrimiento; y que podían imaginarse utilizando la aplicación en su vida diaria. No obstante, los usuarios demandaron una mayor velocidad y sesiones de evaluación de menor duración (Martínez-Cagigal et al., 2019a). A pesar de la creciente popularidad de los *smartphones*, existen pocas aproximaciones previas en la literatura BCI dedicadas a controlar alguna funcionalidad de estos dispositivos. De hecho, más del 56 % del tiempo dedicado a los *smartphones* se emplea en socializar (Ipsos MORI and Google, 2017). Hasta donde alcanza nuestro conocimiento, no existe ningún estudio previo que permita utilizar estos dispositivos a un alto nivel o controlar alguna red social mediante un sistema BCI. Existen estudios que permiten llamar a contactos (Campbell et al., 2010; Chi et al., 2012; Wang, 2010), aceptar llamadas entrantes (Katona et al., 2014), abrir aplicaciones preinstaladas (Elsawy and Eldawlatly, 2015), deletrear palabras (Elsawy et al., 2017; Obeidat et al., 2017) o controlar un juego sencillo (Wu et al., 2014). Sin embargo, ninguna de ellas ha sido testeada con personas con grave discapacidad, e incluso los resultados obtenidos para los HS son inferiores a los alcanzados en nuestro estudio. Por esta razón, nuestra aproximación puede considerarse como uno de los estudios precursores en el control de dispositivos móviles mediante sistemas BCI, orientados a su uso por parte de personas con grave discapacidad.

C.6 Conclusiones

A raíz de los resultados obtenidos en la presente Tesis Doctoral se pueden extraer las siguientes conclusiones:

1) Un sistema BCI basado en potenciales P300 práctico debe implementar una etapa de asincronía para evitar selecciones de comandos involuntarias. Este proceso es esencial para proveer al usuario de un control completo sobre el sistema y evitar la dependencia en supervisores externos, favoreciendo la

autonomía de los usuarios objetivo.

2) La integración de la etapa de detección asíncrona mejora significativamente el rendimiento de los usuarios en aplicaciones BCI de asistencia. La cantidad de errores cometidos se reduce drásticamente.

3) Las señales EEG de los usuarios mientras atienden a las estimulaciones visuales en paradigmas *oddball* son significativamente más complejas e irregulares que cuando las ignoran. Estas diferencias permiten monitorizar la atención del usuario mediante técnicas basadas en entropía.

4) El conjunto de canales óptimo depende en gran medida en el individuo, reflejando una alta variabilidad inter-sujeto. Por tanto, una optimización personalizada para cada usuario es beneficiosa para el rendimiento general del sistema BCI, y constituye una práctica recomendada si se dispone del tiempo necesario tras la calibración.

5) Las metaheurísticas multiobjetivo discretas son apropiadas para encontrar conjuntos de canales óptimos en función del número de electrodos a usar, y superan significativamente a la indicación general de emplear 8 canales en los sistemas P300-BCI. Una combinación balanceada entre técnicas deterministas y estocásticas (e.g., DFGA) favorece la convergencia.

6) Los rendimientos de las personas con graves discapacidades motoras son significativamente más bajos que los obtenidos en sujetos de control. Por tanto, los sistemas BCI de asistencia deben evaluarse con usuarios objetivo para asegurar su viabilidad en un entorno real.

7) Las aplicaciones BCI para controlar un navegador web y redes sociales móviles han demostrado su utilidad como asistencia a personas con grave discapacidad. Su integración en la vida diaria de personas dependientes es viable.

8) Las opiniones de los participantes reflejan una satisfacción general con las aplicaciones de asistencia propuestas. Los pacientes se imaginan usando los sistemas P300-BCI en su vida diaria en un futuro cercano.

Bibliography

Acqualagna, L., Blankertz, B., 2013. Gaze-independent BCI-spelling using rapid serial visual presentation (RSVP). Clinical Neurophysiology 124 (5), 901–908.

Alcaide-Aguirre, R. E., Warschausky, S. A., Brown, D., Aref, A., Huggins, J. E., 2017. Asynchronous brain–computer interface for cognitive assessment in people with cerebral palsy. Journal of Neural Engineering 14 (066001), 1–10.

Allison Brendan Z., F. J. N. C., 2012. BCIs that use steady-state visual evoked potentials or slow cortical potentials. In: Wolpaw, J., Wolpaw, E. W. (Eds.), Brain-computer interfaces: principles and practice. Oxford University Press, New York, pp. 45–78.

Aloise, F., Schettini, F., Aricò, P., Leotta, F., Salinari, S., Mattia, D., Babiloni, F., Cincotti, F., 2011. P300-based brain–computer interface for environmental control: an asynchronous approach. J. Neural Eng. 8 (2), 25025.

Aref, A., Huggins, J., 2012. The P300-certainty algorithm: improving accuracy by withholding erroneous selections. In: EEG and Clinical Neuroscience Society Conference 2012.

Arican, M., Polat, K., 2020. Binary particle swarm optimization (BPSO) based channel selection in the EEG signals and its application to speller systems. Journal of Artificial Intelligence and Systems 2 (1), 27–37.

Aydin, E. A., Bay, O. F., Guler, I., 2018. P300-Based Asynchronous Brain Computer Interface for Environmental Control System. IEEE Journal of Biomedical and Health Informatics 22 (3), 653–663.

Barten, P. G. J., 1999. Contrast sensitivity of the human eye and its effects on image quality. Vol. 21. Spie optical engineering press Bellingham, WA.

Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological) 57 (1), 289–300.

Bensch, M., Karim, A. a., Mellinger, J., Hinterberger, T., Tangermann, M., Bogdan, M., Rosenstiel, W., Birbaumer, N., 2007. Nessi: An EEG-Controlled Web Browser for Severely Paralyzed Patients. Computational intelligence and neuroscience 2007, 71863.

Berger, H., 1929. Über das Elektroenzephalogramm des Menschen. Archiv für psychiatrie und nervenkrankheiten 87 (1), 527–570.

Bishop, C. M., 2006. Pattern recognition and machine learning. springer.

Blankertz, B., Lemm, S., Treder, M., Haufe, S., Müller, K. R., 2011. Single-trial analysis and classification of ERP components - A tutorial. NeuroImage 56 (2), 814–825.

Blankertz, B., Müller, K.-R., Krusienski, D. J., Schalk, G., Wolpaw, J. R., Schlögl, A., Pfurtscheller, G., Millán, J. D. R., Schröder, M., Birbaumer, N., 2006. The BCI competition III: Validating alternative approaches to actual BCI problems. IEEE Transactions on Neural Systems and Rehabilitation Engineering 14 (2), 153–159.

Bojorges-Valdez, E., Yanez-Suarez, O., 2018. Association between EEG spectral power dynamics and event related potential amplitude on a P300 speller. Biomedical Physics and Engineering Express 4 (2).

Bozorg-Haddad, O., Solgi, M., A. Loaiciga, H., 2017. Meta-Heuristic and Evolutionary Algorithms for Engineering Optimization.

Breitwieser, C., Pokorny, C., Müller-Putz, G. R., 2016. A hybrid three-class brain-computer interface system utilizing SSSEPs and transient ERPs. Journal of Neural Engineering 13 (6).

Bronzino, J. D., Peterson, D. R., 2014. Biomedical engineering fundamentals. CRC press.

Brunner, C., Birbaumer, N., Blankertz, B., Guger, C., Kübler, A., Mattia, D., Millán, J. d. R., Miralles, F., Nijholt, A., Opisso, E., Ramsey, N., Salomon, P., Müller-Putz, G. R., 2015. BNCI Horizon 2020: towards a roadmap for the BCI community. Brain-Computer Interfaces 2 (1), 1–10.

Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nature reviews neuroscience 10 (3), 186.

Campbell, A., Choudhury, T., Hu, S., Lu, H., Mukerjee, M. K., Rabbi, M., Raizada, R. D., 2010. NeuroPhone: brain-mobile phone interface using a wireless EEG headset. Proceedings of the second ACM SIGCOMM workshop on Networking, systems, and applications on mobile handhelds - MobiHeld '10, 3–8.

Chatrian, G. E., Lettich, E., Nelson, P. L., 1985. Ten Percent Electrode System for Topographic Studies of Spontaneous and Evoked EEG Activities. American Journal of EEG Technology 25 (2), 83–92.

Chaurasiya, R. K., Londhe, N. D., Ghosh, S., 2017. Multi-objective binary DE algorithm for optimizing the performance of Devanagari script-based P300 speller. Biocybernetics and Biomedical Engineering 37 (3), 422–431.

Chi, Y. M., Wang, Y. T., Wang, Y., Maier, C., Jung, T. P., Cauwenberghs, G., 2012. Dry and noncontact EEG sensors for mobile brain-computer interfaces. IEEE Transactions on Neural Systems and Rehabilitation Engineering 20 (2), 228–235.

Coello, C. a., Reyes-Sierra, M., 2006. Multi-Objective Particle Swarm Optimizers: A Survey of the State-of-the-Art. International Journal of Computational Intelligence Research 2 (3), 287–308.

Coello, C. A. C., Lamont, G. B., 2004. Applications of multi-objective evolutionary algorithms. Vol. 1. World Scientific.

Cohen, M. X., 2014. Analyzing neural time series data: theory and practice. MIT press.

Colwell, K. A., Ryan, D. B., Throckmorton, C. S., Sellers, E. W., Collins, L. M., 2014. Channel selection methods for the P300 Speller. Journal of Neuroscience Methods 232, 6–15.

Compston, A., Coles, A., 2008. Multiple sclerosis. Lancet 372 (9648), 1502–1517.

Costa, M., Goldberder, A. L., Peng, C.-K., 2002. Multiscale Entropy Analysis of Complex Physiologic Time Series. Physical Review Letters 89 (6), 1–4.

Costa, M., Goldberger, A. L., Peng, C. K., 2005. Multiscale entropy analysis of biological signals. Physical Review E - Statistical, Nonlinear, and Soft Matter Physics 71 (2), 1–18.

Curry, D. M., Dagli, C. H., 2014. Computational complexity measures for many-objective optimization problems. Procedia Computer Science 36 (C), 185–191.

Deb, K., 2005. Multi-Objective Optimization.

Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation 6 (2), 182–197.

Delatycki, M. B., Williamson, R., Forrest, S. M., 2000. Friedreich ataxia: An overview. Journal of Medical Genetics 37 (1), 1–8.

Donchin, E., Ritter, W., McCallum, W. C., 1978. Cognitive psychophysiology: The endogenous components of the ERP. In: Event-related brain potentials in man. pp. 349–411.

Donnan, G. A., Fisher, M., Macleod, M., Davis, S. M., 2008. Stroke. The Lancet 371 (9624), 1612–1623.

Dorigo, M., Birattari, M., Stutzle, T., 2006. Ant colony optimization. IEEE Computational Intelligence Magazine 1 (4), 28–39.

Eiben, A. E., Smith, J. E., 2003. Introduction to Evolutionary Computing. Vol. 53. Berlin.

Elsawy, A. S., Eldawlatly, S., 2015. P300-based applications for interacting with smart mobile devices. In: 7th Annual International IEEE EMBS Conference on Neural Engineering. pp. 166–169.

Elsawy, A. S., Eldawlatly, S., Taher, M., Aly, G. M., 2017. MindEdit: a P300-based text editor for mobile devices. Computers in Biology and Medicine 80 (August 2016), 97–106.

Farcomeni, A., 2008. A review of modern multiple hypothesis testing, with particular attention to the false discovery proportion. Statistical Methods in Medical Research 17 (4), 347–388.

Farina, D., Jensen, W., Akay, M., 2013. Introduction to Neural Engineering for Motor Rehabilitation. John Wiley & Sons.

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

Ferrari, M., Quaresima, V., 2012. A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. NeuroImage 63 (2), 921–935.

Ferree, T. C., Luu, P., Russell, G. S., Tucker, D. M., 2001. Scalp electrode impedance, infection risk, and EEG data quality. Clinical neurophysiology 112 (3), 536–544.

Fodor, J. A., 1983. The modularity of mind. MIT press.

Fuchs, E. C., Zivkovic, A. R., Cunningham, M. O., Middleton, S., LeBeau, F. E., Bannerman, D. M., Rozov, A., Whittington, M. A., Traub, R. D., Rawlins, J. N. P., Monyer, H., 2007. Recruitment of Parvalbumin-Positive Interneurons Determines Hippocampal Function and Associated Behavior. Neuron 53 (4), 591–604.

Gonzalez, A., Nambu, I., Hokari, H., Iwahashi, M., Wada, Y., 2013. Towards the classification of single-trial event-related potentials using adapted wavelets and particle swarm optimization. Proceedings - 2013 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2013, 3089–3094.

Gonzalez, A., Nambu, I., Hokari, H., Wada, Y., 2014. EEG Channel Selection Using Particle Swarm Optimization for the Classification of Auditory Event-Related Potentials. Scientific World Journal 2014, 350270.

Guger, C., Vaughan, T., Allison, B., 2014. Brain-Computer Interface Research: A State-of-the-Art Summary 3.

He, S., Zhang, R., Wang, Q., Chen, Y., Yang, T., Feng, Z., Zhang, Y., Shao, M., Li, Y., 2017. A P300-Based Threshold-Free Brain Switch and Its Application in Wheelchair Control. IEEE Transactions on Neural Systems and Rehabilitation Engineering 25 (6), 715–725.

Hochberg Leigh R., A. K. D., 2012. BCI users and their needs. In: Wolpaw, J., Wolpaw, E. W. (Eds.), Brain-computer interfaces: principles and practice. Oxford University Press, New York, pp. 45–64.

Holland, J. H., 1992. Genetic algorithms. Scientific american 267 (1), 66–73.

Huettel, S. A., McCarthy, G., 2004. What is odd in the oddball task? Prefrontal cortex is activated by dynamic changes in response strategy. Neuropsychologia 42 (3), 379–386.

Humeau-Heurtier, A., 2015. The multiscale entropy algorithm and its variants: A review. Entropy 17 (5), 3110–3123.

Ikeda, A., Lüders, H. O., Burgess, R. C., Shibasaki, H., 1992. Movement-related potentials recorded from supplementary motor area and primary motor area: role of supplementary motor area in voluntary movements. Brain 115 (4), 1017–1043.

Ipsos MORI, Google, 2017. Something for everyone: why the growth of mobile apps is good news for brands. Tech. rep.

Jahanshahi, M., Hallett, M., 2003. The Bereitschaftspotential: Movement-Related Cortical Potentials. Springer, New York.

Jin, J., Allison, B. Z., Brunner, C., Wang, B., Wang, X., Zhang, J., Neuper, C., Pfurtscheller, G., 2010. P300 Chinese input system based on Bayesian LDA. Biomedizinische Technik 55 (1), 5–18.

Jobson, J. D., 1991. Applied multivariate data analysis: volume I: Regression and Experimental Design. Springer.

Jr, I. F., Yang, X.-s., Fister, I., Brest, J., 2013. A Brief Review of Nature-Inspired Algorithms for Optimization. arXiv preprint arXiv:1307.4186 80 (3).

Karim, A. A., Hinterberger, T., Richter, J., Mellinger, J., Neumann, N., Flor, H., Kubler, A., Birbaumer, N., 2006. Neural Internet: Web Surfing with Brain Potentials for the Completely Paralyzed. Neurorehabilitation and Neural Repair 20 (4), 508–515.

Katona, J., Peter, D., Ujbanyi, T., Kovari, A., 2014. Control of incoming calls by a Windows Phone based Brain Computer Interface. 15th IEEE International Symposium on Computational Intelligence and Informatics (CINTI 2014), 121–125.

Kee, C.-Y., Ponnambalam, S., Loo, C.-K., 2015. Multi-objective genetic algorithm as channel selection method for P300 and motor imagery data set. Neurocomputing 161, 120–131.

Kemp, S., 2018. Digital in 2018. Tech. rep.

Kennedy, J., Eberhart, R., 1995. Particle swarm optimization. Neural Networks, 1995. Proceedings., IEEE International Conference on 4, 1942—-1948 vol.4.

Kisley, M. A., Cornwell, Z. M., 2006. Gamma and beta neural activity evoked during a sensory gating paradigm: Effects of auditory, somatosensory and cross-modal stimulation. Clinical Neurophysiology 117 (11), 2549–2563.

Kleber, B., Birbaumer, N., 2005. Direct brain communication: Neuroelectric and metabolic approaches at Tübingen. Cognitive Processing 6 (1), 65–74.

Krusienski, D. J., Sellers, E. W., Cabestaing, F., Bayoudh, S., McFarland, D. J., Vaughan, T. M., Wolpaw, J. R., 2006. A comparison of classification techniques for the P300 Speller. J. Neural Eng. 3 (4), 299–305.

Krusienski, D. J., Sellers, E. W., McFarland, D. J., Vaughan, T. M., Wolpaw, J. R., 2008. Toward enhanced P300 speller performance. J. Neurosci. Methods 167 (1), 15–21.

Krusienski Dean J., M. D. J. P. J. C., 2012. BCI Signal Processing: Feature Extraction. In: Wolpaw, J., Wolpaw, E. W. (Eds.), Brain-computer interfaces: principles and practice. Oxford University Press, New York, pp. 45–64.

Kübler, A., Kotchoubey, B., Kaiser, J., Birbaumer, N., Wolpaw, J. R., 2001. Brain-computer communication: Unlocking the locked in. Psychol. Bull. 127 (3), 358–375.

Kübler, A., Nijboer, F., Kleih, S., 2020. Hearing the needs of clinical users. In: Handbook of Clinical Neurology. Vol. 168. Elsevier, pp. 353–368.

Li, Y., Pan, J., Wang, F., Yu, Z., 2013. A hybrid BCI system combining P300 and SSVEP and its application to wheelchair control. IEEE Transactions on Biomedical Engineering 60 (11), 3156–3166.

Lu, M.-k., Arai, N., Tsai, C.-H., Ziemann, U., 2012. Movement Related Cortical Potentials of Cued Versus Self-Initiated Movements : Double Dissociated Modulation by Dorsal Premotor Cortex Versus Supplementary Motor Area rTMS. Human Brain Mapping 33, 824–839.

Luck, S. J., 2014. An introduction to the event-related potential technique. MIT press.

Ma, Z., Qiu, T., 2018. Quasi-periodic fluctuation in Donchin's speller signals and its potential use for asynchronous control. Biomedizinische Technik 63 (2), 105–112.

Martínez-Cagigal, V., Gomez-Pilar, J., Álvarez, D., Hornero, R., 2017. An asynchronous P300-based Brain–Computer Interface web browser for severely disabled people. IEEE Transactions on Neural Systems and Rehabilitation Engineering 25 (8), 1332–1342.

Martínez-Cagigal, V., Hornero, R., 2017a. A Binary Bees Algorithm for P300-Based Brain-Computer Interfaces Channel Selection. In: Advances in Computational Intelligence. IWANN 2017. Lecture Notes in Computer Science, 1st Edition. Springer International Publishing AG, Cádiz, Spain, pp. 453–463.

Martínez-Cagigal, V., Hornero, R., 2017b. P300-based Brain-Computer Interface channel selection using swarm intelligence. Revista Iberoamericana de Automática e Informática Industrial 14 (4), 372–383.

Martínez-Cagigal, V., Santamaría-Vázquez, E., Gomez-Pilar, J., Hornero, R., 2019a. Towards an accessible use of smartphone-based social networks through brain-computer interfaces. Expert Systems with Applications 120, 155–166.

Martínez-Cagigal, V., Santamaría Vázquez, E., Hornero, R., 2018. A Novel Hybrid Swarm Algorithm for P300-Based BCI Channel Selection. In: Proceedings of the World Congress on Medical Physics & Biomedical Engineering (Vol. 3) (IUPESM 2018). Springer, Prague, pp. 41–45.

Martínez-Cagigal, V., Santamaría-Vázquez, E., Hornero, R., 2019b. Asynchronous Control of P300-Based Brain–Computer Interfaces Using Sample Entropy. Entropy 21 (3), 230.

Martínez-Cagigal, V., Santamaría-Vázquez, E., Hornero, R., 2020. Brain–Computer Interface Channel Selection Optimization using Meta-heuristics and Evolutionary Algorithms. Applied Soft Computing R2.

McCane, L. M., Sellers, E. W., Mcfarland, D. J., Mak, J. N., Carmack, C. S., Zeitlin, D., Wolpaw, J. R., Vaughan, T. M., 2014. Brain-computer interface (BCI) evaluation in people with amyotrophic lateral sclerosis. Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration 15 (3-4), 207–215.

McDonald, J. W., Sadowsky, C., 2002. Spinal-cord injury. The Lancet 359 (9304), 417–425.

McIntosh, A. R., 1999. Mapping cognition to the brain through neural interactions. Memory 7 (5-6), 523–548.

Mercuri, E., Muntoni, F., 2013. Muscular dystrophies. Lancet (London, England) 381 (9869), 845–60.

Miller, E. K., Freedman, D. J., Wallis, J. D., 2002. The prefrontal cortex: categories, concepts and cognition. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 357 (1424), 1123–1136.

Miller, L. E., Hatsopoulos, N., 2012. Neuronal Activity in motor cortex and related areas. In: Wolpaw, J., Wolpaw, E. W. (Eds.), Brain-computer interfaces: principles and practice. Oxford University Press, New York, pp. 15–44.

Mugler, E. M., Ruf, C. a., Halder, S., Bensch, M., Kübler, A., 2010. Design and implementation of a P300-based brain-computer interface for controlling an internet browser. IEEE Transactions on Neural Systems and Rehabilitation Engineering 18 (6), 599–609.

Narsky, I., Porter, F. C., 2013. Statistical analysis techniques in particle physics: fits, density estimation and supervised learning. John Wiley & Sons.

Nicolas-Alonso, L. F., Gomez-Gil, J., 2012. Brain computer interfaces, a review. Sensors 12 (2), 1211–1279.

NINDS, 2013. Cerebral Palsy: Hope Through Research. Tech. rep., National Institutes of Health (NIH), Bethesda (Maryland).

Nunez, P., Srinivasan, R., Westdorp, A., Wijesinghe, R., Tucker, D., Silberstein, R., Cadusch, P., 1997. EEG coherency I: statistics, reference electrode, volume conduction, Laplacians, cortical imaging, and interpretation at multiple scales. Electroencephalography and Clinical Neurophysiology 103 (5), 499–515.

Nunez, P. L., 2012. Electric and magnetic fields produced by the brain. In: Wolpaw, J., Wolpaw, E. W. (Eds.), Brain-computer interfaces: principles and practice. Oxford University Press, New York, pp. 45–64.

Obeidat, Q., Campbell, T., Kong, J., 2017. Spelling with a small mobile Brain-Computer Interface in a moving wheelchair. IEEE Transactions on Neural Systems and Rehabilitation Engineering 4320 (c), 1–1.

Ofner, P., Schwarz, A., Pereira, J., Müller-Putz, G. R., 2017. Upper limb movements can be decoded from the time-domain of low-frequency EEG. PLoS ONE 12 (8), 1–24.

Ogawa, S., Lee, T. M., Kay, A. R., Tank, D. W., 1990. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. In: Proceedings of the National Academy of Sciences. Vol. 87. pp. 9868–9872.

Panicker, R. C., Puthusserypady, S., Pryana, A. P., Sun, Y., 2010. Asynchronous P300 BCI: SSVEP-based control state detection. European Signal Processing Conference 58 (6), 934–938.

Panicker, R. C., Puthusserypady, S., Sun, Y., 2011. An asynchronous P300 BCI with SSVEP-based control state detection. IEEE Transactions on Biomedical Engineering 58 (6), 1781–1788.

Pastor, M. A., Artieda, J., Arbizu, J., Valencia, M., Masdeu, J. C., 2003. Human cerebral activation during steady-state visual-evoked responses. The Journal of neuroscience : the official journal of the Society for Neuroscience 23 (37), 11621–7.

Pelillo, M., 2013. Similarity-Based Pattern Analysis and Recognition. Springer, London.

Perseh, B., Sharafat, A. R., jun 2012. An Efficient P300-based BCI Using Wavelet Features and IBPSO-based Channel Selection. Journal of Medical Signals and Sensors 2 (3), 128–143.

Pfurtscheller, G., Lopes da Silva, F. H., 1999. Event-related EEG/MEG synchronization and desynchronization: basic principles. Clinical Neurophysiology 110 (11), 1842–1857.

Pfurtscheller, G., McFarland, D. J., 2012. BCIs that user sensorimotor rhythms. In: Wolpaw, J., Wolpaw, E. W. (Eds.), Brain-computer interfaces: principles and practice. Oxford University Press, New York, pp. 45–78.

Pham, D. T., Ghanbarzadeh, A., Koç, E., Otri, S., Rahim, S., Zaidi, M., 2006. The Bees Algorithm - A Novel Tool for Complex Optimisation Problems. Intelligent Production Machines and Systems - 2nd I*PROMS Virtual International Conference, 454–459.

Pickles, J., 2013. An introduction to the physiology of hearing. Brill.

Picton, T. W., 1992. The P300 wave of the human event-related potential. Journal of Clinical Neurophysiology 9 (4), 456–479.

Pinegger, A., Faller, J., Halder, S., Wriessnegger, S. C., Müller-Putz, G. R., 2015. Control or non-control state: that is the question! An asynchronous visual P300-based BCI approach. Journal of Neural Engineering 12 (1), 014001.

Priemer, R., 1990. Introductory signal processing. Vol. 6. World Scientific Publishing Company.

Reilly, B., 2013. Mu Rhythm. In: Volkmar, F. R. (Ed.), Encyclopedia of Autism Spectrum Disorders. Springer New York, New York, NY, pp. 1940–1941.

Reyes-Sierra, M., Coello, C. A., 2005. Improving PSO-Based Multi-objective Optimization Using Crowding, Mutation and E-Dominance. Lecture Notes in Computer Science 3410, 505–519.

Richman, J. S., Moorman, J. R., 2000. Physiological time-series analysis using approximate entropy and sample entropy. American Journal of Physiology-Heart and Circulatory Physiology 278 (6), H2039–H2049.

Robbins, K., min Su, K., Hairston, W. D., 2018. An 18-subject EEG data collection using a visual-oddball task, designed for benchmarking algorithms and headset performance comparisons. Data in Brief 16, 227–230.

Santamaría-Vázquez, E., Martínez-Cagigal, V., Gomez-Pilar, J., Hornero, R., 2019. Asynchronous Control of ERP-Based BCI Spellers Using Steady-State Visual Evoked Potentials Elicited by Peripheral Stimuli. IEEE Transactions on Neural Systems and Rehabilitation Engineering 27 (9), 1883–1892.

Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N., Wolpaw, J. R., 2004. BCI2000: A general-purpose brain-computer interface (BCI) system. IEEE Transactions on Biomedical Engineering 51 (6), 1034–1043.

Sellers Eric W., A. Y. D. E., 2012. BCIs that use P300 event-related potentials. In: Wolpaw, J., Wolpaw, E. W. (Eds.), Brain-computer interfaces: principles and practice. Oxford University Press, New York, pp. 215–226.

Sirvent Blasco, J., Iáñez, E., Úbeda, A., Azorín, J., 2012. Visual evoked potential-based brain–machine interface applications to assist disabled people. Expert Systems with Applications 39 (9), 7908–7918.

Society for Neuroscience, 2017. BrainFacts.
URL `https://www.brainfacts.org/`

Sörnmo, L., Laguna, P., 2005. Bioelectrical signal processing in cardiac and neurological applications. Academic Press.

Speier, W., Arnold, C., Pouratian, N., 2013. Evaluating true BCI communication rate through mutual information and language models. PLoS ONE 8 (10).

Srinivasan, R., 2012. Acquiring brain signals from outside the brain. In: Wolpaw, J., Wolpaw, E. W. (Eds.), Brain-computer interfaces: principles and practice. Oxford University Press, New York, pp. 105–122.

Stam, C. J., 2004. Functional connectivity patterns of human magnetoencephalographic recordings: A 'small-world' network? Neuroscience Letters 355 (1-2), 25–28.

Standring, S., 2016. Gray's anatomy: the anatomical basis of clinical practice, 41st Edition. Elsevier Ltd, London.

Storn, R., Price, K., 1997. Differential Evolution - A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. Journal of Global Optimization 11 (4), 341–359.

Tang, J., Liu, Y., Jiang, J., Yu, Y., Hu, D., Zhou, Z., 2018. Toward Brain-Actuated Mobile Platform. International Journal of Human-Computer Interaction 00 (00), 1–12.

Tangermann, M., Müller, K. R., Aertsen, A., Birbaumer, N., Braun, C., Brunner, C., Leeb, R., Mehring, C., Miller, K. J., Müller-Putz, G. R., Nolte, G., Pfurtscheller, G., Preissl, H., Schalk, G., Schlögl, A., Vidaurre, C., Waldert, S., Blankertz, B., 2012. Review of the BCI competition IV. Frontiers in Neuroscience 6 (JULY), 1–31.

Tatum IV, W. O., 2014. Handbook of EEG Interpretation, 2nd Edition. New York : Demos Medical.

Toma, K., Matsuoka, T., Immisch, I., Mima, T., Waldvogel, D., Koshy, B., Hanakawa, T., Shill, H., Hallett, M., 2002. Generators of Movement-Related Cortical Potentials : fMRI-Constrained EEG Dipole Source Analysis. NeuroImage 17, 161–173.

Townsend, G., LaPallo, B. K., Boulay, C. B., Krusienski, D. J., Frye, G. E., Hauser, C. K., Schwartz, N. E., Vaughan, T. M., Wolpaw, J. R., Sellers, E. W., 2010. A novel P300-based brain-computer interface stimulus presentation paradigm: moving beyond rows and columns. Clin. Neurophysiol. 121 (7), 1109–1120.

Treder, M. S., Schmidt, N. M., Blankertz, B., 2011. Gaze-independent brain-computer interfaces based on covert attention and feature attention. J Neural Eng 8 (6), 66003.

Tripathi, P. K., Bandyopadhyay, S., Pal, S. K., 2007. Multi-Objective Particle Swarm Optimization with time variant inertia and acceleration coefficients. Information Sciences 177 (22), 5033–5049.

Vidal, J. J., 1973a. Toward direct brain-computer communication. Annual review of Biophysics and Bioengineering 2 (1), 157–180.

Vidal, J. J., 1973b. Toward Direct Brain-Computer Communication. Annual Review of Biophysics and Bioengineering 2 (1), 157–180.

Vidal, J. J., 1977. Real-Time Detection of Brain Events in EEG. In: Proceedings of the IEEE. Vol. 65. pp. 633–641.

Vilic, A., 2013. The AVI SSVEP Dataset.
URL `http://www.setzner.com/avi-ssvep-dataset/`

Volosyak, I., Cecotti, H., Gräser, A., 2009. Optimal visual stimuli on LCD screens for SSVEP based brain-computer interfaces. 2009 4th International IEEE/EMBS Conference on Neural Engineering, NER '09, 447–450.

Wang, X.-J., 2010. Neurophysiological and Computational Principles of Cortical Rhythms in Cognition. Physiological Reviews 90 (3), 1195–1268.

Wang, Y.-T., Wang, Y., Jung, T.-P., 2011. A cell-phone-based brain-computer interface for communication in daily life. Journal of neural engineering 8, 025018.

Witten, I. H., Frank, E., 2011. Data Mining: Practical Machine Learning Tools and Techniques, 3rd Edition. Morgan Kaufmann.

Wolpaw, J., Wolpaw, E. W., 2012a. Brain-computer interfaces: principles and practice. OUP USA.

Wolpaw, J. R., Birbaumer, N., Heetderks, W. J., McFarland, D. J., Peckham, P. H., Schalk, G., Donchin, E., Quatrano, L. A., Robinson, C. J., Vaughan, T. M., 2000. Brain-computer interface technology: a review of the first international meeting. IEEE Transactions on Rehabilitation Engineering 8 (2), 164–173.

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., Vaughan, T. M., 2002. Brain-computer interfaces for communication and control. Clinical Neurophysiology 113 (6), 767–91.

Wolpaw, J. R., Wolpaw, E. W., 2012b. Brain–computer interfaces: something new under the sun. In: Wolpaw, J., Wolpaw, E. W. (Eds.), Brain-computer interfaces: principles and practice. Oxford University Press, New York, pp. 3–14.

Wolpaw, J. R., Wolpaw, E. W., 2012c. Signals reflecting brain metabolic activity. In: Wolpaw, J., Wolpaw, E. W. (Eds.), Brain-computer interfaces: principles and practice. Oxford University Press, New York, pp. 65–77.

Wolpaw Jonathan R., W. E. W., 2012. The future of BCIs: Meeting the expectations. In: Wolpaw, J., Wolpaw, E. W. (Eds.), Brain-computer interfaces: principles and practice. Oxford University Press, New York, pp. 45–64.

Wu, G., Xie, Z., Wang, X., 2014. Development of a mind-controlled Android racing game using a brain computer interface (BCI). 2014 4th IEEE International Conference on Information Science and Technology, 652–655.

Yang, X.-S., 2014. Nature-Inspired Optimization Algorithms, 1st Edition. Elsevier Inc.

Yu, T., Li, Y., Long, J., Gu, Z., 2012. Surfing the internet with a BCI mouse. Journal of Neural Engineering 9 (3), 036012.

Yu, Y., Zhou, Z., Jiang, J., Yin, E., Liu, K., Wang, J., Liu, Y., Hu, D., 2017. Toward a Hybrid BCI: Self-Paced Operation of a P300-based Speller by Merging a Motor Imagery-Based "Brain Switch" into a P300 Spelling Approach. International Journal of Human-Computer Interaction 33 (8), 623–632.

Zhang, H., Guan, C., Wang, C., 2008. Asynchronous P300-based brain-computer interfaces: a computational approach with statistical models. IEEE Transactions on Biomedical Engineering 55 (6), 1754–63.

Zitzler, E., Laumanns, M., Thiele, L., 2001. SPEA2: Improving the Strength Pareto Evolutionary Algorithm. Tech. rep., ETH Zurich, Zurich.

Zweig, M. H., Campbell, G., 1993. Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. Clinical Chemistry 39 (4), 561–577.

Index

T

W

For decades, mankind has fantasized about the possibility of controlling devices with our minds. Despite that there is still a long way to go to achieve that goal, recent progresses in neuroscience took a step forward and contributed to the development of the first brain–computer interface (BCI) systems. Through the analysis of the electroencephalographic signals, BCIs are able to decode users' intentions into application commands. Nevertheless, they currently do not provide the required reliability to take the leap from laboratories to real environments. The present Doctoral Thesis is focused on the development of novel signal processing techniques and assistive applications that contribute to provide a real use of P300-based BCIs by motor-disabled people. In particular, asynchrony and channel optimization problems are addressed, and two assistive applications (web browsing and mobile social networks) are proposed. In view of the results, it is believed that this compendium of publications will contribute the BCI literature to move toward a real of these systems by the motor-disabled, aiming at improving their personal autonomy and quality of life.

UVa

Doctoral Thesis
Compendium of publications
International Mention