User Adapted Brain-Computer Interface

Federica Turi

▶ **To cite this version:**

Federica Turi. User Adapted Brain-Computer Interface. Signal and Image processing. Université Côte d'Azur, 2020. English. NNT : . tel-03149221v1

HAL Id: tel-03149221

https://inria.hal.science/tel-03149221v1

Submitted on 22 Feb 2021 (v1), last revised 23 Mar 2021 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ CÔTE D'AZUR

ÉCOLE DOCTORALE
SCIENCES ET TECHNOLOGIES DE L'INFORMATION ET DE LA COMMUNICATION

THÈSE DE DOCTORAT

Interface cerveau-ordinateur adaptée à l'utilisateur

Federica TURI

Inria Sophia Antipolis – Méditerranée, Athena Project Team

Présentée en vue de l'obtention du grade de docteur en
Sciences et Technologies de l'Information et de la Communication
d'Université Côte d'Azur

Dirigée par: Maureen Clerc

Soutenue le : 23/09/2020

Devant le jury, composé de :

Peter Desain, Professeur,
Donders Centre for Cognition - Rapporteur

Fabien Lotte, Directeur de Recherches,
Inria Bordeaux Sud-Ouest - Rapporteur

Maureen Clerc, Directrice de Recherches,
Inria Sophia Antipolis - Directrice de thèse

Sylvain Chevallier, Maître de conférences,
LISV - Examinateur

Théodore Papadopoulo, Directeur de Recherches,
Inria Sophia Antipolis - Examinateur

Athena

Inria

Resumé

Les interfaces cerveau-ordinateur (Brain-Computer Interface ou BCI) permettent la communication entre l'utilisateur et la machine, grâce à la traduction de l'activité cérébrale en commandes qui servent à contrôler différents dispositifs. De nombreuses limitations empêchent la diffusion des systèmes BCI dans des applications réelles, telles que la phase de calibration qui résulte de la variabilité entre sessions et entre sujets. Cette phase est fondamentale car elle permet de régler les paramètres nécessaires pour le bon fonctionnement du système, mais elle est considérée beaucoup trop longue et fatigante pour le sujet.

L'objectif de cette thèse est de surmonter ces limites par de nouvelles méthodes basées sur l'amélioration ou le remplacement de la phase de calibration traditionnelle, en proposant le développement d'une BCI centrée sur l'utilisateur.

D'abord, nous proposons deux systèmes BCI adaptés au sujet. Le premier concerne un clavier virtuel basé sur des potentiels évoqués modulés par code-modulated Visual Evoked Potential (c-VEP) où la phase de calibration traditionnelle est remplacée par une phase dans laquelle les paramètres de stimulation sont réglés de manière adaptée au sujet. Le deuxième concerne le développement d'un système basé sur l'imagination mentale (MI-BCI) pour un sujet à déficience motrice sévère (tétraplégie), dans le cadre d'une compétition internationale BCI en direct.

Ensuite, nous proposons une auto-calibration pour le c-VEP BCI qui utilise les propriétés fondamentales de la réponse VEP pour prédire les mots à l'aide d'un dictionnaire, éliminant la phase de calibration.

Les méthodes proposées ont montré des résultats prometteurs et ouvrent de nouvelles perspectives pour la diffusion des systèmes BCI plus adaptables à l'utilisateur.

Mots-clés : Electroencéphalographie (EEG), Interface cerveau-ordinateur, Traitement du signal, Potentiels évoqués visuels

Abstract

Brain-Computer Interface (BCI) allows communication between an user and a machine, by converting the user's brain activity into commands that control external devices. Many limitations prevent the diffusion of BCI systems in real applications, such as the calibration phase that is a consequence of the issue of variability across sessions and among users. The calibration phase is fundamental because it allows to set the main parameters to extract the relevant information from the electroencephalograpy (EEG) signal of the subject, but it is considered time consuming and tedious for the user.

The objective of this thesis is to overcome these limitations by novel methods based on the improvement or even replacement of the traditional calibration phase, proposing the development of a user-centered BCI system.

Firstly, we present a design to develop an adaptive BCI system for two different applications. The former deals with a code-modulated Visual Evoked Potential (c-VEP) speller where an adaptive parameter setting phase is proposed to replace the standard calibration phase. The latter application concerns the development of a Mental Imagery (MI) BCI for a disabled user, characterized by a long user-centered multi-stage training phase, in the context of a live BCI competition.

Secondly, we propose an auto-calibration c-VEP BCI system exploiting the language information. In our model the fundamental properties that characterize the VEP response are used to predict the full word using a dictionary, eliminating the traditional calibration phase.

The proposed methods showed promising results and open new perspectives to the diffusion of BCI systems more adaptable to the user.

Keywords: Brain-Computer Interface, Electroencephalograpy (EEG), Signal Processing, Visual Evoked Potentials

Acknowledgements

These years have been a special journey for me and many people have been part of it.

I would like to thank Maureen Clerc, my supervisor, that gave me this unique opportunity and followed me during these years. Thanks to all the members of the Athena team, that shared these years with me and thanks to Rachide Deriche and Theo Papadoupolo that received me in the team. Thanks to Romain Lacroix, who gave a fundamental contribution in the implementation of my BCI software. Thanks to all the volunteers that participated to my experiments, thank you because you dedicated for free your time to my research, giving a very important contribution to the BCI community. I would like to thank all the people of the Nitro team that faced this challenge and shared with me this unique experience, with a special thank to Karine Leclerc that accepted with enthusiasm this challenge and to Amandine Audino, Pierre Giacalone and Paul-Emmanuel Ponsenard.

Thanks to the professors of the Polytech Nice Sophia, that gave me the possibility of teaching for four years. This was a very formative experience for me.

A special thanks to my family, my parents and my sisters, Eleonora and Valeria, that always support my choices with love, even if far away.

Finally, I would like to thank Luca who supports, motivates, encourages, inspires me and makes special every single day of my life.

Contents

List of Figures

List of Tables

List of Abbreviations

BCI	Brain-Computer Interface
CCA	Canonical Correlation Analysis
c-VEP	code-modulated Visual Evoked Potential
EEG	Electroencephalography
ECG	Electrocardiography
ECoG	Electrooculography
ERD	Event Related Desynchronization
ERS	Event Related Synchronization
EMG	Electromyography
FFT	Fast Fourier Transform
GUI	Graphic User Interface
ITR	Information Transfer Rate
LDA	Linear Discriminant Analysis
MI	Mental Imagery
ROC curve	Receiver Operating Characteristic curve
SSVEP	Steady State Visual Evoked Potential
SMRs	Senosry Motor Rhythms
SNR	Signal-to-Noise Ratio
VEP	Visual Evoked Potential

Part I

Introduction

Introduction

Context

The main purpose of Brain-Computer Interfaces (BCIs) is to restore communication skills for people with severe motor disabilities, converting electrophysiological input from the user's brain into an output that controls external devices. The objective of a BCI system is to interpret the user's intention by extracting features from his/her brain activity. Indeed, there are specific patterns in the EEG signal which can be used as control signals in a BCI system [Townsend et al., 2010; Yuan and He, 2014]. In this thesis we focused on Visual Evoked Potentials (VEPs), that are brain activity patterns generated by external stimulation, and on sensorimotor rhythms (SMRs), which are specific oscillations generated by mental strategies without external stimulation.

BCI applications are widespread in many research fields [Abdulkader et al., 2015], with the hope that this technology could really improve life of people with severe motor disabilities. There are many areas of applications, such as communication and entertainment.

In this work a special attention is given to two types of applications: VEP BCI spellers, that allow linguistic communication by word spelling, and the SMR-BCI for a severely disabled user in the context of a BCI-game competition.

Objective

Many limitations [Lotte et al., 2015] prevent the diffusion of BCI systems in real applications. These limitations are challenges for BCI research to find solutions to improve the diffusion of the system outside the research laboratory, in order to really improve the quality of daily-life of impaired people.

Among these limitations there is the tedious calibration phase. This phase is fundamental in a BCI system because it allows the system to set the preprocessing and calibration parameters to extract the relevant information from the EEG signal, in terms of spatial filters and classifiers. This phase is fundamental but is also really annoying, long and tedious for the subject and considered time consuming [Lotte, 2015].

The calibration is a consequence of the problem of the inter-subject and inter-session variability. Many works underlined the importance of a calibration-free BCI, proposing different strategies to reduce the calibration time, such as approaches based on transfer learning methods [Jayaram et al., 2016; Kindermans et al., 2012b; Lotte, 2015].

Hence, to face the inter-session and inter-subject variability issues the research is oriented on the development of user-centered BCI systems, in order to develop an user-friendly system able to adapt to any condition of the subject independently of the situation.

The goal of this thesis is to explore novel methods that replace or avoid the traditional calibration phase with a component in the system that allows it to learn and detect the specific brain patterns of each user, independently of the sessions, opening the path for the development of an user-centered efficient BCI system for a code-Visual Evoked Potential (c-VEP) BCI speller and a Motor-Imagery (MI) BCI for an impaired user.

In comparison to other BCI paradigms, the Visual Evoked Potentials BCIs yield a high communication speed [Bin et al., 2011; Nakanishi et al., 2017]. Nevertheless, the c-VEP BCI system is still limited in diffusion in the BCI research community and applications. Thus, the work is focused on the development of an user-centered approach to improve the usability of a c-VEP system, with the aim of opening new research topics for the diffusion of this type of system in real applications.

In a c-VEP BCI speller the calibration phase is fundamental to improve the robustness of the system but directly impacts its usability, indeed it is considered time-consuming and tedious for the user and prevents the diffusion of this system in daily life. To tackle this limitation we proposed two approaches. The former approach deals with the improvement of the traditional calibration phase, in terms of speed, usefulness and comfort of the

user. The latter proposes a unsupervised classifier, in order to avoid the traditional calibration phase.

Moreover the usability challenge was faced for a MI-BCI proposing the design of an user-centered multi-task training for an impaired user in order to develop an effective system, even in stressful scenario, such as a live BCI competition.

Contributions

This review is focused on the development of novel strategies to develop an user-centered system, with the following contributions.

Development of adaptive BCI systems We proposed the design of a BCI system in two different contexts. Firstly, we designed an adaptive c-VEP BCI speller in order to tackle the inter-session and inter-subject variability. We proposed an adaptive parameter setting phase that replaces the tedious standard calibration phase, in order to define the more pleasant and efficient stimulus for each subject across sessions. Secondly, we deployed a MI-BCI system for a disabled user in the context of a BCI-game competition. We present the long multi-stage training necessary to obtain a good performance and highlight that in this type of system the user has a central role. Moreover we analyze the impact of the user's state in a real scenario.

Auto-calibration of c-VEP BCI by word prediction We propose an unsupervised classifier for a c-VEP BCI system exploiting only the language information. The proposed model uses the fundamental property that characterizes the c-VEP response to predict the full word using a dictionary, eliminating the calibration phase. The proposed method shows promising results to develop a calibration-free c-VEP BCI and also opens a new perspective to the diffusion of BCI systems more adaptable to the user.

Overview of the Thesis

Chapter 1. Fundamentals of Brain-Computer Interface

In the first chapter, we present the fundamental concepts of a BCI system. Firstly, we detail the BCI system and its components. Then, a focus is made on the EEG signal, detailing the acquisition method and the specific EEG artifacts. After, the most common brain activity patterns used to design EEG-BCI systems are presented, describing how the EEG activity can be applied to BCI systems. Finally, a general overview of the main fields of BCI applications concludes this chapter.

Chapter 2. Endogenous end Exogenous BCI Systems and Limitations

In the second chapter, we present two BCI applications: the ERP and VEP BCI spellers and SMR-based BCI. The BCI speller applications are presented describing their pipeline and referring to the state-of-the-art. A special attention is given to the importance of the GUI in a speller.

Then, the SMR-based BCI system is, briefly, introduced with a focus on disabled user. We analyze and discuss the limitations of these applications and the possible solutions.

Chapter 3. Design and Flashing Strategy for BCI Spellers

In Chapter 3, we present two pilot studies to investigate and implement different stimulation strategies in VEP BCI spellers. The aim of this chapter is to show the impact of some stimulus parameters on the performance of a BCI speller.

The former deals with the implementation of a flashing strategy in a P300 system in order to improve the discriminability between flashed characters and to evaluate their influence on the performance of the speller. We tested the system on healthy subjects and we discuss the results obtained.

The latter is based on the design of a c-VEP BCI speller, aiming to provide some new features for a BCI speller GUI and software starting from scratch. We tested and discuss different experimental setups to prove the effectiveness of the developed system.

Chapter 4. Adaptive BCI Systems

In Chapter 4, two different works concerning the realization of a user-centered system are presented.

The first consists in the development of an adaptive c-VEP BCI speller. We propose a possible solution to the inter-session and inter-subject variability by an adaptive parameter setting c-VEP BCI system. In the proposed system the traditional calibration phase is replaced by a shorter adaptive setting phase, in order to define the more pleasant and efficient stimulus for each subject across sessions. We present the experimental setup and the results reached after conducting an experiment involving nine healthy subjects.

The second work concerns the realization of a MI-BCI system for a motor-impaired user in the context of a BCI-game competition. In particular, we present the long training necessary to obtain a performing system. Then, we discuss about our participation to the competition, with a focus on the factors that influenced the pilot's performance.

Chapter 5. Auto-Calibration of c-VEP BCI by Word Prediction

In Chapter 5, we propose a proof of concept for an auto-calibration c-VEP BCI speller. We introduce the novel method and provide results reached on a public subject dataset, in order to prove its feasibility. We discuss those results, dealing also with current state-of-art of BCI spellers.

Chapter 6. Conclusion and Perspectives

Finally, we conclude this manuscript summarizing the main contributions of the thesis, with a general discussion on the achieved results, limitations and future perspectives of the proposed methods in BCI field.

Part II

Background

1 Fundamentals of Brain-Computer Interface

1.1 Brain-Computer Interface system

A Brain-Computer Interface (BCI) is a hardware and software communication system that converts input from the user's brain into an output that controls external devices, such as a computer, a robotic prosthesis or a speech synthesizer. It is defined as a system that measures Central Nervous System (CNS) activity and converts it into artificial output that replaces, restores or improves natural CNS output and thereby changes the ongoing interactions between the CNS and external environments [Wolpaw and Wolpaw, 2012]. The main goal of BCI research is to provide communication capabilities to severely disabled people who are totally paralyzed or "locked in" by neurological neuromuscular disorders, such as amyotrophic lateral sclerosis (ALS), brain stem stroke, muscular dystrophies or spinal cord injury [Birbaumer and Cohen, 2007; Daly and Wolpaw, 2008; Van Dokkum et al., 2015].

The objective of BCI is a translation algorithm that converts brain input from the user into an output that controls external devices. A BCI must function as an adaptive closed-loop control system, indeed BCI operation depends on effective interaction between two adaptive controllers. The former is the user who encodes his/her commands in the electro-physiological input provided to the BCI, the latter is the BCI system that recognizes the commands contained in the input and expresses them in device control and gives a real-time feedback to the user, who can optimize his/her brain activity in order to obtain the desired output [Daly and Wolpaw, 2008].

A BCI system acts like a control system between the user and the external device. The input in this system is the brain activity of the user and the output is the commands on the external device, that can provide a specific stimulation

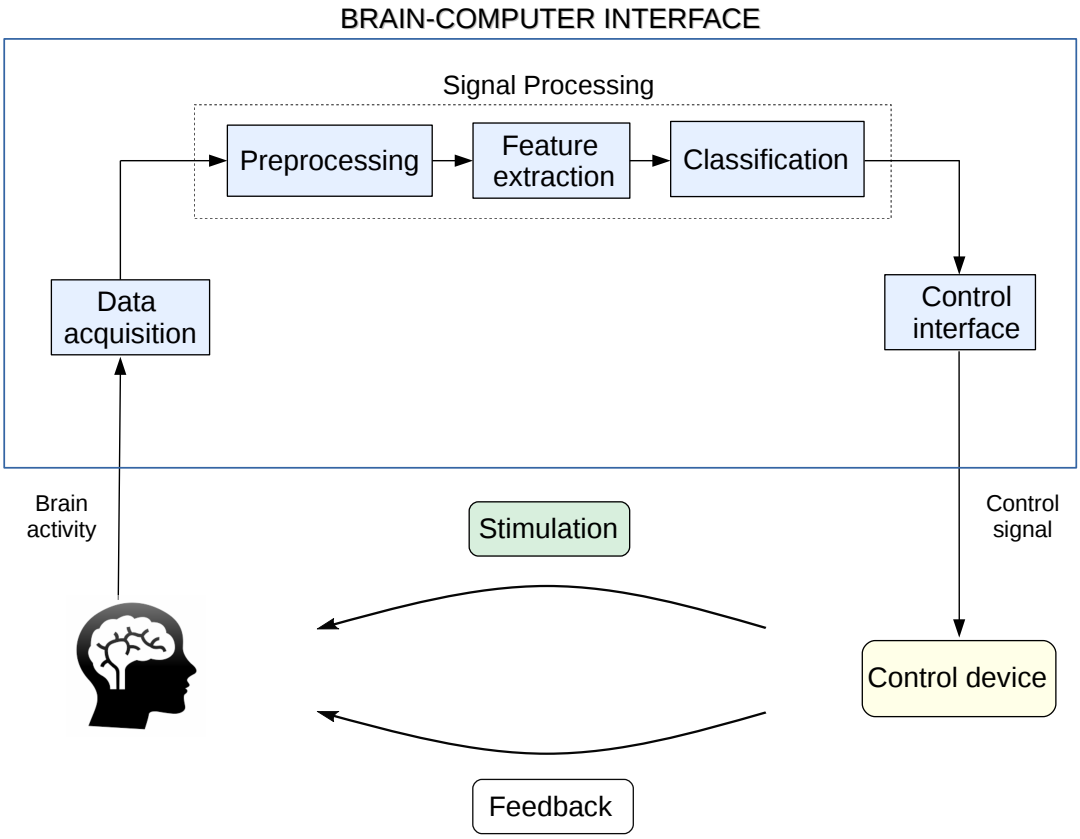


FIGURE 1.1 – Architecture of a BCI system consisting of three steps: data acquisition, signal processing that includes preprocessing of the signal, extraction of the features of interest in the signal, classification of the features, and a control interface to interact with an external device.

to the user. To allow this translation from input to output, three steps are necessary: signal acquisition, signal processing and control interface [Wolpaw et al., 2002; Nicolas-Alonso and Gomez-Gil, 2012]. The signal processing phase includes preprocessing, feature extraction and classification phases. These are outlined in Figure 1.1.

Data acquisition

The objective of the *data acquisition* stage is to record and digitize the brain activity. The brain activity is used as input source in a BCI system to gather information on user's intentions. There are several techniques to measure brain activity [Wolpaw and Wolpaw, 2012; Nicolas-Alonso and Gomez-Gil, 2012]. One of the most common methods is the electroencephalogram (EEG) [Wolpaw et al., 2002], that is also the method on which we focused this review.

We will introduce the EEG acquisition system and the EEG signal in the next Section 1.2.1.

Signal processing

The *signal processing* stage aims to detect and extract the user's intentions from the raw brain activity. It consists in two main steps: feature extraction and classification. The objective of the feature extraction is to apply methods to extract from the brain signal the relevant features specifically related to the user's intentions. First of all it is fundamental, but not mandatory, to deploy a preprocessing step in which some specific methods are applied to enhance the raw cerebral signals by temporal filtering, noise reduction and artifact removing. In general, temporal or spatial filtering techniques are applied to increase the Signal-to-Noise Ratio (SNR) of the EEG. Then specific algorithms are applied to extract the features, that usually are put into a vector, called *feature vector*.

For BCI based on the EEG there are three sources of information from which the features can be extracted: spatial information, spectral information and temporal information [Lotte, 2014]. It depends on the type of BCI to use the right sources of information. For example, BCI based on event related potential (e.g. P300 speller) or based on code Visual Evoked Potential (e.g. c-VEP BCI) need temporal and spatial information to extract features.

The second step of the signal processing consists in classifying the specific signal features found during the previous step, to obtain commands that carry out the user's intent. The classification consists basically in the assignment of the signal features to specific categories, called *classes*. Each class generates a specific output or commands for the system. The most commonly used classification algorithms in the BCI community are for example Artificial Neural Networks (ANNs), which are well established in BCI research and have numerous successful applications [Yang et al., 2012], Linear Discriminant Analysis (LDA) classifier [Hoffmann et al., 2008] and Support Vector Machine (SVM) classifiers. For an extensive review of BCI classification methods, we refer the reader to [Lotte et al., 2007; Lotte et al., 2018].

Control interface

The *control interface* stage translates the classified signals into commands that are sent to control an external device, such as a speller, a mouse cursor, a motorized wheelchair or a neuroprosthetic arm. The action of the device gives the feedback to the user, closing the control loop.

1.2 Signal acquisition in BCIs

There are two main types of data acquisition techniques to record brain activity: non-invasive and invasive techniques. The most common invasive technique is electrocorticography (ECoG) [Schalk and Leuthardt, 2011], that records the brain activity directly from the surface of the brain. The advantage of this method is that the SNR of the signal is high because the electrodes are directly in contact with the brain. The disadvantage is that the electrodes have to be placed by surgery, so the implantation is really invasive for the user and not very practical, and it also involves ethical problems.

Instead a non-invasive technique does not require a surgery to measure the brain activity, so the risks for the user are reduced. There are various assessment techniques for different types of measured signals such as Positron Emission Tomography (PET), functional Magnetic Resonance Imaging (fMRI) [Weiskopf et al., 2003], functional Near-Infrared Spectroscopy (fNIRS) [Coyle et al., 2004] [Naseer and Hong, 2015], magnetoencephalography (MEG) [Mellinger et al., 2007], and electroencephalography (EEG).

1.2.1 EEG signal

Among non-invasive methodologies the most widespread in BCI applications is the EEG [Wolpaw et al., 2000], that allows to record the brain activity of different brain areas through a set of sensors placed on the scalp of the user. The EEG technique is widely used in BCI applications for different reasons: the EEG equipment is of relatively low cost, the set-up is easy, it involves little risk to the users, it has an high portability and high temporal resolution, with respect to others non-invasive techniques such as fMRI [Van Gerven et al., 2009]. For these reasons the EEG technique is widely used in the BCI

field, it became an useful tool in laboratories of research and development of BCI systems, where the preliminary studies involved healthy users. Nevertheless, the downside of the EEG is that between the brain and the sensors that acquire the signals there are many layers, that increase the noise on the signal, reduce the SNR, and they can limit the informative content.

Brain rhythms

The EEG measures the electrical activity of the cerebral cortex. The neural activity detectable by the EEG is the summation of the excitatory and inhibitory postsynaptic potentials of tens of thousands synchronous neurons [Britton et al., 2016]. The EEG waves may be classified according to their frequencies, amplitude, shape and distribution over the scalp [Kumar and Bhuvaneswari, 2012]. The brain wave frequencies belong to a range from 0.1 Hz to 100 Hz and are associated with different mental states, with different states of consciousness or pathological disorders [Turi, 2015]. The frequency bands categorized in EEG are delta (δ), theta (θ), alpha (α), beta (β), and gamma (γ). The main characteristics of these bands are detailed below.

The *delta band* includes frequencies between $0.1 - 4$ Hz. The appearance of delta waves is normal in neonatal and infants EEG and during sleep stages in adult EEG. The presence of delta rhythm in the waking adult EEG indicates cerebral injury or severe cerebral disease. Delta waves appear also during coma.

The *theta band* is between $4 - 8$ Hz. Larger presence of theta activity is normal in infancy and childhood. The EEG of an adult subject contains only a small amount of theta frequencies, mostly observed in states of drowsiness. As for the delta waves, the presence of theta oscillations in the waking adult is abnormal and is symptom of various pathological problems.

The *alpha band* includes frequency between $8 - 13$ Hz. Alpha rhythm appears during wakefulness (in a normal adult person), mostly over the occipital regions of the head. Alpha waves are best observed with eyes closed and in subjects who are in a condition of physical relaxation. They are attenuated when the patient returns in a situation of alertness, especially visual. Hence these rhythms may be useful to measure mental effort, because the increasing

of mental effort induces reduction of alpha activity [Venables and Fairclough, 2009]. The alpha band is called mu (μ) for movement tasks.

The *beta band* covers a range of frequencies between $13 - 30$ Hz. The beta activity is usually recorded in the frontal and central regions of the brain. It can be registered in almost every healthy adult, because it is associated with motor activity and with active thinking, such as problem solving and decision making.

The *gamma band* has a frequency range between $30 - 100$ Hz. The gamma rhythms are not spontaneous, but usually they are evoked by sensory stimulation. It is difficult to detect gamma activity with the scalp EEG, but easier with invasive technique, such as the ECoG.

1.2.2 EEG recording

EEG recording is obtained by electrodes placed on the scalp. The aim of the electrode is to pick up bioelectric signals and transform them to electric current. An electrode acts as a transducer which basically consists of an electrical conductor in contact with aqueous ionic solution, such as a special gel. Among electrodes for scalp EEG there are dry electrodes, that do not require gel, and gel electrodes. They allow to measure the electrical potentials generated by the activity of tens of thousands of neurons within the brain, through a non-invasive way. Scalp EEG is also easily corrupted by both physiological and non-physiological artifacts. Physiological artifacts include sweat, chewing, eye-blinks and scalp muscle contraction. Non-physiological artifacts include power-line noise and EEG cable motion. Dry electrodes are really recent in the EEG field, they are more practical to use because they do not need the application of gel and allow to reduce this time-consuming procedure when performing EEG measurements. However they often introduce new problems, like signal quality and comfort [Guger et al., 2012].

Electrode placement montage

In order to obtain EEG recordings there are several standardized systems for electrode placement on the skull, useful to obtain consistency between laboratories. The most common electrode placement strategy is based on

the standard 10–20 system [Klem, 1999], shown in Figure 1.2. It is the conventional electrode setting of the International EEG Federation, in which anatomic landmarks are defined to correctly positioning the electrodes on the scalp. The positions are determined using as reference points the nasion (delve at the top of the nose), the inion (bony lump at the base of the skull on the mid-line at the back of the head) and the pre-auricular points anterior to the ear. The distances between adjacent electrodes are either 10 % or 20 % of the total front–back or right–left distance of the skull, this explains why the system is called "10–20". To record a more detailed EEG with higher spatial resolution, extra electrodes are added using the 10 % division, which fills the intermediate space between the electrodes of the 10–20 system with additional electrodes, as detailed in Figure 1.2.

The standard electrode position nomenclature uses letters and numbers. The letters indicate the particular lobe or brain area: *F* frontal lobe, *P* parietal lobe, *O* occipital lobe, *T* temporal lobe, *C* refers to central electrodes, *FP* refers to the centropolar electrodes, which are placed above the eyes and *A* refers to the ear lobe. The letters are followed by numbers (Figure 1.2). Even numbers represent the right side of the head; odd numbers refer to the left side. The middle is called "*Z*". The numbering system starts at the mid-line (*Z*) and progresses laterally.

It is possible to obtain unipolar (or monopolar) and bipolar derivations from these electrodes. A bipolar derivation measures the potential difference between pair of electrodes. Therefore, bipolar derivations give an indication of the potential gradient between two cerebral areas. In unipolar setup, the potential difference is measured relative to the same electrode for all derivations. This reference electrode is usually placed on the earlobe, nose, chin or neck. It is possible to obtain an *average reference* montage by subtracting from each channel the average activity from all derivations.

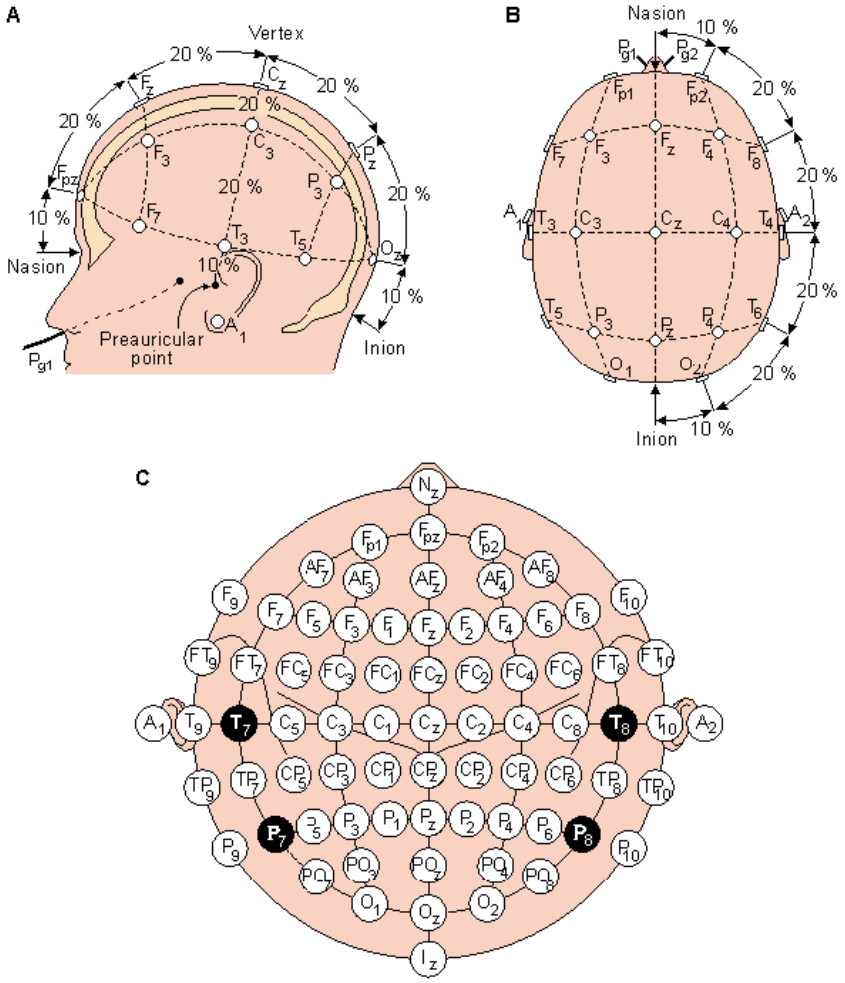


FIGURE 1.2 – Electrode locations of International 10–20 system for EEG recording [He et al., 2013]

1.2.3 EEG artifacts

Artifacts refer to any recorded electrical activity which has not cerebral origin. It can be a physiologic or an extra-physiologic artifact, if it is generated from the patient or from the environment. It is fundamental to detect and remove those artifacts, in order to avoid misinterpretation of EEG data [Jiang et al., 2019; Islam et al., 2016].

Physiologic artifacts

Physiologic artifacts are generated from the patient, from sites other than the brain. The principal physiologic artifacts are related to cardiac, ocular and muscular artifacts [Libenson, 2012; Tatum IV, 2018].

— **Cardiac artifacts** The heart produces two different types of artifacts: electrical and mechanical ones. They are both correlated with cardiac cycle and are easily identified as they are synchronized with ECG: the electrical cardiac artifact is actually the ECG signal recorded from scalp electrodes and in EEG the QRS complexes are visible. Mechanical cardiac artifacts (pulse artifacts) are related to the circulatory pulse and can be considered as electrode artifacts as well: they occur when an electrode is placed over a vessel, which pulses and produces a periodic slow wave in EEG signals. Usually this type of artifact is visible over the frontal and temporal regions. It can be removed by repositioning the interested electrode a little farther from the vessel.

— **Ocular artifacts** The most visible ocular artifact is the eye blinking artifact: it produces a bifrontal potential with very typical morphology. This artifact is caused by the eyes' electrical dipoles, oriented along the corneal-retinal axis: when the eyes open, close or move, the dipole creates a moving electrical field that causes the artifact on the EEG signal. Also lateral eye movements create artifacts with typical morphology, different from that of eye blinks. Slow lateral movements are very common during drowsiness and generate an artifact that usually occurs in an intermittent way, accompanied by slowing of the alpha rhythm.

— **Muscle artifacts** A movement during the recording of the EEG could produce artifacts due to both a movement effect on some electrodes

and to an electrical field generated by muscles (electromyography activity). This EMG signal overlaps and obscures the EEG because it has greater amplitude, and it usually occurs in frontal and temporal regions.

Extra-physiologic artifacts

Extra-physiologic artifacts are generated from outside the patient, mainly from medical instrumentation and the environment. Extra-physiologic artifacts include electrode artifacts and line noise [Libenson, 2012; Tatum IV, 2018].

— **Electrode artifacts** appear usually in the form of few brief waveforms, limited to one electrode on the scalp. The most common of these artifacts is the one due to electrode pops: it reflects the ability of the electrode-skin interface to act as a capacitor and store electrical charge across the gel that fixes the electrode in its position on the scalp. When the release of this charge occurs, there is a fast change in impedance and a potential appears in all the channels that include the electrode. Another type of electrode artifact is the one caused by poor contact with the skin: this type of artifact is represented in the EEG as sharp or slow waves with varying morphology and amplitude, due to instability in the impedance.

— **Line noise** is caused by the alternating current present in the electrical power supply (50 or 60 Hz). It is usually present in many channels.

1.3 Brain activity pattern in BCIs

The objective of a BCI system is to interpret the user's intention by extracting features from his/her brain activity. There are specific patterns in the EEG signal which can be used as control signals in a BCI system and according to their nature, BCI systems can be classified as either exogenous or endogenous [Nicolas-Alonso and Gomez-Gil, 2012].

In particular if the brain activity patterns used in BCIs as control signals are generated by external stimulation, such as in Visual Evoked Potentials (VEPs) and Event-Related Potentials (P300), the BCI system is classified as

exogenous. Otherwise, if the control signals are generated by the detection of specific oscillations not associated with external stimulation, like sensorimotor rhythms and slow cortical potentials, the system is classified as endogenous.

1.3.1 Visual Evoked Potentials

A Visual Evoked Potential (VEP) is an electrical potential elicited by visual stimuli, which can be recorded in the visual cortex by surface EEG [Townsend et al., 2010]. The VEP has long been used in clinical neuroscience to measure the functional integrity of the visual system from retina to the visual cortex of the brain, via the optic nerves. Today, it is mostly used as a functional test that can help to detect eye diseases, but it is also used in the BCI field as input modality to develop BCI systems.

In a VEP-BCI system a target is coded by a stimulus sequence which evokes a specific VEP recorded in the EEG of the subject. Depending on the specific stimulus modulation design used, current VEP based BCIs can be distinguished into BCI systems using time-modulated VEP (t-VEP) [Lee et al., 2008], frequency-modulated VEP (f-VEP) [Middendorf et al., 2000; Cheng et al., 2002; Bin et al., 2009b] and BCI systems using pseudo-random code-modulated VEP (c-VEP) [Bin et al., 2009a].

1.3.2 P300 Evoked Potentials

Event-Related Potentials (ERPs) are very low intensity voltage differences generated in response to specific events or stimuli [Sur and Sinha, 2009]. ERPs appears in the EEG as positive or negative deflections from the baseline activity. The one most utilized in BCI field is the P300, a positive detection that is recorded on the EEG 300 ms after the specific stimulus. The "oddball" paradigm [Donchin and Coles, 1988] is the most utilized paradigm to elicit the P300. In this paradigm different stimuli are presented in a series such that one of them occurs relatively infrequently. The subject has to focus on the infrequent target stimulus and not on the frequently presented or standard stimulus. Whenever the target stimulus occurs, a P300 is elicited in the brain of the subject.

1.3.3 Sensorimotor Rhythms

Sensorimotor Rhythms (SMRs) are specific rhythms recorded on the EEG exclusively over sensorimotor areas (i.e. μ-rhythm and β-rhythm). SMRs occurs during mental tasks, such as motor imagery, that block or increase the power of the ongoing EEG signal. Hence SMRs do not need external stimulation to be generated, because the specific mental strategies modulate the activation and deactivation of the sensorimotor cortex [Yuan and He, 2014]. The increases or decreases of power in given frequency bands and spatial location can be used to identify the mental task that caused this change in the brain rhythm. The decrease in brain rhythm is called an event-related desynchronization (ERD), while an increase is called event-related synchronization (ERS) [Pfurtscheller and Da Silva, 1999]. For instance during motor imagery task such as, for example, imagination of left-hand, right-hand, or foot movement, a ERD is observed in both μ and β bands before the movement, while after the end of the movement a ERS occurs in the β band.

1.3.4 Slow Cortical Potentials

The Slow Cortical Potentials (SCPs) are slow voltage changes recorded in the EEG in a low frequency band, around 1 Hz. After a substantial training, an user can learn to self-regulate his/her SCPs and then can utilize this skill to control, for example, a Language Support Program [Kübler et al., 2001]. However SCPs-based BCI is not often used because an user needs long training phase to gain a full control of his/her SCPs, and not all users are able to gain the full control.

1.4 BCI Applications

BCI applications are widespread in many research fields [Abdulkader et al., 2015], with the hope that this technology can really improve life of people with severe motor disabilities. Roughly, there are three main areas of applications: communication, motor restoration and entertainment. BCIs for communication control are BCI systems designed to provide a new method

of communication for impaired people to interact with the external environment. Current BCI systems are applied for word spelling (BCI speller) and for environmental control (e.g. lights, television, computer) [Cincotti et al., 2008]. Many BCI systems are proposed for restoring motor function. The state-of-art proposes different applications, but the most common are BCIs designed to control a neuroprothesic device or a mobile robot [Barbosa et al., 2010]. Also BCI systems for locomotion can be included in this research field, locomotion-oriented BCI systems allow people with severe disabilities to autonomously drive a wheelchair [Tanaka et al., 2005] or control a exoskeleton [Benabid et al., 2019]. BCIs for communication and for motor restoration are applied as assistive applications for disabled patients. However, with the increasing interest in entertainment and gaming BCI applications, the BCI use is extended also to healthy subjects. A BCI can be introduced in video games to provide new information from the brain activity, such as emotional conditions that can be used as input into the game.

Moreover, many other applications have been developed for healthy users in different fields such as smart-home, workplaces and transportation, in order to improve the human's daily life [Abdulkader et al., 2015]. Many of these applications involve passive BCIs [Zander and Kothe, 2011], in which the system passively estimates the user's mental state to control the BCI application. Thus, the user can be concentrated on his/her specific daily tasks without needing to focus on the BCI system.

In the next section some specific applications, such as speller BCIs and gaming BCIs for disabled subjects will be detailed and discussed, analyzing the actual state-of-art, limitations and solutions.

1.5 Conclusions

In this chapter were presented the fundamental concepts of a BCI systems. First of all, the main components of a BCI system were introduced: the data acquisition, the signal processing and the control interface. Then, a focus was made on the EEG signal, detailing the acquisition method and the specific EEG artifacts. After, the most common brain activity patterns used to design specific EEG-BCI systems were presented, describing how this activity can

be applied to BCI systems. Finally, a general overview of the main fields of BCI applications concluded this chapter.

Part III

State of The Art in BCI Applications

2 Endogenous end Exogenous BCI Systems and Limitations

In the previous chapter we described the main part of a BCI system and the brain activity applied in BCI applications. In particular we made the distinction between endogenous end exogenous BCI system.

In the first part of this chapter we deeply present the enxogenous BCI speller that requires external visual stimuli, such as the P300 speller and VEP speller, that are the main topics of this thesis.

In the second part we briefly introduce the MI-BCI with a specific attention on the application for motor-impaired user.

2.1 The BCI spellers

One of the main objective of BCIs is to restore the linguistic communication by word spelling, for this reason BCI speller is one of the most widely used BCI application, since the first spelling application that was developed by [Farwell and Donchin, 1988]. A BCI speller allows the user to restore the communication with his/her external environment through a virtual keyboard, that presents a visual stimulation. The VEP stimulation is perceived by the eye of the user and the signal is transmitted via the optic nerve to the visual cortex, that can be located in the occipital lobe, as shown in Figure 2.1. Instead, in the case of the P300 the potential is observed mainly in the central and parietal regions of the cerebral cortex [Linden, 2005]. In this way the user can select a specific character thanks to the brain activity recorded and processed by the BCI system.

The design of the Graphical User Interface (GUI) [Rezeika et al., 2018] is one of the most important factors to develop a performing BCI speller. The GUI

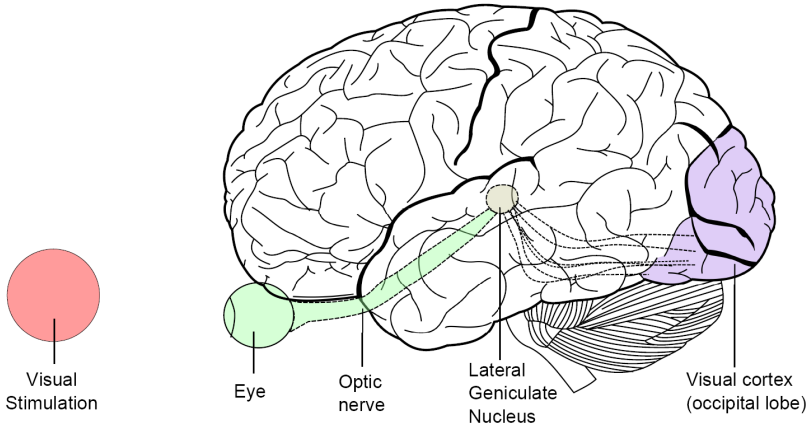


FIGURE 2.1 – A visual stimulation is sent to the eye, with a specific stimulus modulation that depends of the VEP BCI paradigm. This signal is perceived and transmitted via the optic nerve toward the visual cortex, which is located in the occipital lobe. Adapted from [Evain, 2016].

directly affects the performance of the system because it is the front-end of the BCI speller system, it generates stimuli and provides feedback to the user. Hence, the stimulus strategy, that involves the main parameters of the visual stimulation, such as stimulation color, stimulation frequency, keyboard arrangement and many others, is an important factor to investigate and to consider for developing high-performing systems.

The majority of exogenous BCI spellers based on different EEG paradigms, described in Section 1.2.1, are: the P300 speller [Farwell and Donchin, 1988] and the BCI speller that investigates the Visual Evoked Potential (VEP). Among the VEP BCI spellers it is possible to distinguish different systems depending on the specific stimulus modulation design used [Bin et al., 2009a]: the frequency modulated VEP (f-VEP) BCI, such as the steady state evoked potential (SSVEP) BCI [Cecotti, 2010], and BCI systems using pseudo-random code-modulated VEP (c-VEP) [Bin et al., 2011].

The performance of BCI spellers is commonly measured by calculating the accuracy and the Information Transfer Rate (ITR). The accuracy is computed as the ratio of characters well spelled over all the characters spelled.

The ITR [Wolpaw et al., 2002] combines the accuracy and the system speed and returns the amount of information contained in a selection in bits per minute (bits/min or bpm). The objective of this metric is to allow to compare performance for different BCI spellers of the same type. However, recent studies demonstrate that the ITR is not the best way to evaluate the performance of all different types of BCI spellers [Speier et al., 2013], so different metrics are introduced, such as the Output Characters per Minute (OCM) measure, that is computed as the ratio of the total number of characters spelled and the total time spent on spelling these words, usually used in BCI speller that integrate a language model with word completion [Ryan et al., 2010]. Nevertheless, in this thesis we will refer to the ITR parameter for the performance evaluation.

In the following paragraphs the c-VEP speller, the SSVEP speller and the P300 speller will be introduced, focusing on the state-of-the-art of the stimulus modulation that characterizes several types of spellers.

2.1.1 c-VEP BCI speller

In a c-VEP BCI system, all characters flash simultaneously according to a predefined pseudo-random binary sequence circular-shifted by a different time lag [Bin et al., 2011] for each character. The sequences with maximum length, called m-sequences [Golomb, 1982], are generally applied as stimulus sequences in c-VEP BCI. The m-sequence has specific properties of correlation: it is nearly orthogonal to its time shifted sequence and its autocorrelation function is close to a unit impulse function. Therefore, for a given character, the m-sequence generates a VEP response largely dependent on the characteristics of the stimulus code [Bin et al., 2009a]. The VEP response recorded in the electroencephalogram (EEG) of the subject can be used as a template [Spüler et al., 2012]. This template is obtained during a calibration phase at the beginning of each session.

During the calibration phase the user gazes at the reference character C_r and all the characters flash simultaneously for N stimulus cycles. The raw EEG data $\boldsymbol{X_r} \in \mathbb{R}^{n \times c \times s}$ is recorded during $n = 1, ..., N$ stimulus cycles from c channels. The evoked response $\boldsymbol{R} \in \mathbb{R}^{q \times s}$ of the reference character can be obtained averaging the time-windowed EEG data $\boldsymbol{X_r}$ from N stimulus cycles

from q channels. The number of channels q can be the same as c or a subset, if we choose to consider only the channels in which the evoked response is most prominent. To improve the signal-to-noise ratio (SNR) of the system the Canonical Correlation Analysis (CCA) is applied as spatial filter [Spüler et al., 2014] to compute a reference template $T_r(t)$ [Bin et al., 2011]. The goal of CCA is to find the two transformations W_X and W_S which maximize the correlation between the raw EEG data $\boldsymbol{X_r}$ and the expected VEP response $\boldsymbol{S}$, that can be obtained by concatenating N times the evoked response of the reference character.

$$CCA(X,S) = \max_{W_X W_S} \frac{W_X^T \boldsymbol{X}\boldsymbol{S}^T W_S}{\sqrt{W_X^T \boldsymbol{X}\boldsymbol{X}^T W_X} \cdot \sqrt{W_S^T \boldsymbol{S}\boldsymbol{S}^T W_S}} \tag{2.1}$$

Then the spatially filtered EEG data x_n is obtained.

$$x_n = W_X \boldsymbol{X_n} \tag{2.2}$$

The reference template T_r is computed by averaging x_n over N stimulus cycles. The templates T_k of all other K characters in the virtual keyboard are obtained by shifting the reference template T_r by a specific time lag $\tau_k = \tau \cdot k$, where τ is the time lag between the flashing sequences of two consecutive characters and $k = 0, 1, 2, 3, ..., K$ is the index of the corresponding character in the virtual keyboard.

$$T_k(t) = T_r(t - \tau_k) \tag{2.3}$$

The calibration phase is followed by a spelling phase. During the spelling phase the user can gaze at a character of his/her choice, called *target*, for a number of stimulus cycles N. Then, by template matching, it is possible to identify the target that the subject is gazing at (Figure 2.3). The correlation ρ_k (2.4) between the spatially filtered EEG $x = W_X \boldsymbol{X}$, recorded during the online spelling phase, and each template T_k, obtained during the calibration phase, is computed. Hence, the system outputs the detected character k_{target} (2.5) by selecting the one that corresponds to the index of the template with the highest value of correlation.

$$\rho_k = \frac{x^T T_k}{\sqrt{xx^T \cdot T_k T_k^T}} \tag{2.4}$$

$$k_{target} = \arg\max_k \rho_k \tag{2.5}$$

With the explained pipeline a c-VEP BCI can potentially achieve a very high-speed communication level, reaching an average information transfer rate (ITR) of 108 ± 12 bits/min [Bin et al., 2011].

Others classification algorithms are proposed to improve the c-VEP system, such the One-Class Support Vector Machine (OCSVM) proposed by [Spüler et al., 2012]. The results obtained with the OCSVM algorithm showed that it can potentially achieve better performance than the template matching approach in c-VEP BCI, reaching an average ITR of 140 bits/min.

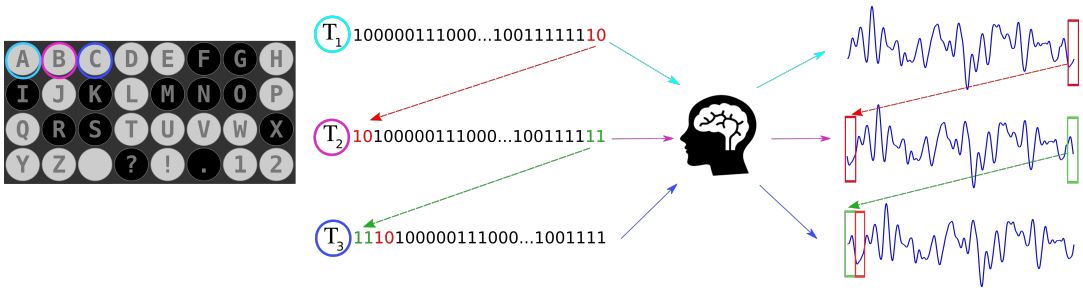


FIGURE 2.2 – An illustration of the generation of c-VEP responses. On the left the m-sequence circular-shifted by a 2-bit lag for each target, on the right the evoked response of each target.

BBVEP BCI speller

A Broad-Band visually evoked potentials (BBVEP) BCI speller [Thielen et al., 2015] is a c-VEP BCI speller, in which each character on the virtual keyboard is flashed following a Gold sequence [Gold, 1967], instead of a circularly shifted m-sequences. A set of Gold sequences is a set of pseudo-random bit sequences generated by cleverly combining two selected m-sequences. This set of pseudo-random bit sequences are characterized by low cross-correlation, indeed a Gold sequence is uncorrelated to any other, including any shifted version of itself. Thus in the case of Gold sequences the set of

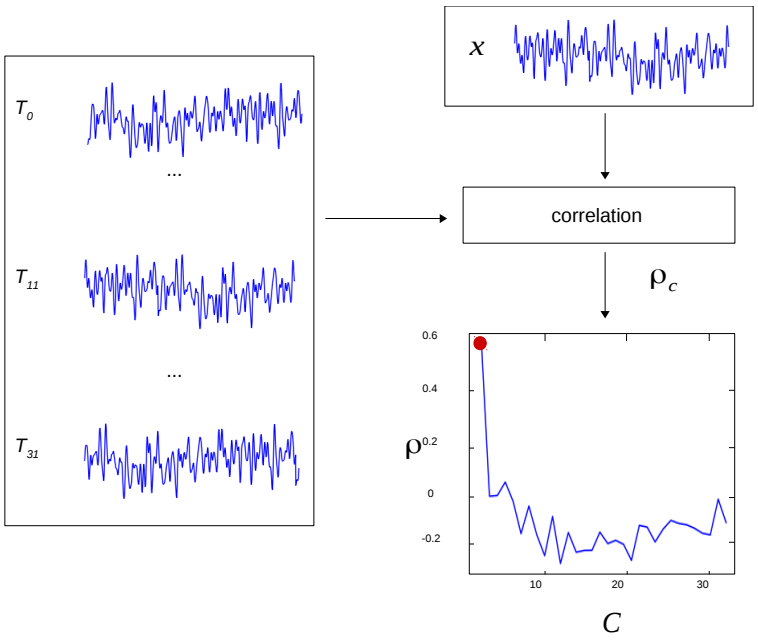


FIGURE 2.3 – Illustration of template matching technique for target identification. The red dot indicates that the character in position 0 (the character "A" in the virtual keyboard of Figure 2.2) is identified as the target character.

stimulation sequences has optimized proprieties of correlation (minimized cross-correlation and maximized auto-correlation) and does not need a synchronization signal. On the other hand for the m-sequences the synchronization is fundamental to discriminate the characters, since they flash all using the same sequence with a different phase-shift.

The authors [Thielen et al., 2015] proposed a template generation strategy using the *reconvolution* method, based on the fact that a Gold sequence is a succession of short and long flashes and that the Broad-Band visually evoked potential brain response at each of these flashes will be the same, so the brain response to each sequence can be find by convolution. The reconvolution method consists in two steps: the *estimation* and the *generation*, as shown in Figure 2.5. In the estimation step the VEP response $x(t)$ is recorded for a trial and decomposed following the structure of a bit-sequence $s_u(t)$, belonging to the U set of Gold codes applied as training, thus the response of the short pulse r_s and the response of the long pulse r_l are obtained. Then in the generation step the template $T(t)$ of an unseen bit-sequence $s_v(t)$, belonging to the V set of Gold codes applied as testing, is predicted by the estimated responses r_s and r_l. In this way it is possible to predict the template for each

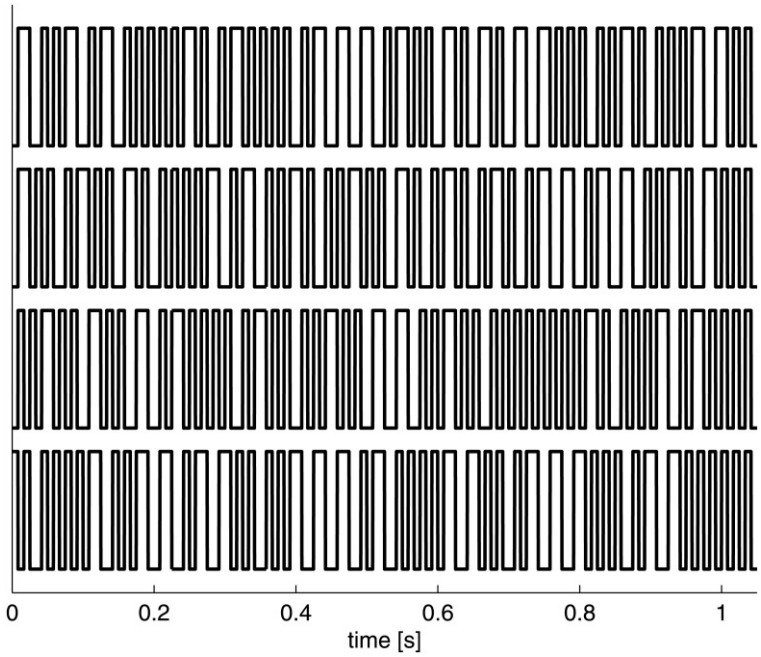


FIGURE 2.4 – Example of set of Golden codes. Adapted from [Thielen et al., 2015].

trial.

Moreover, in order to develop an efficient template matching classifier, they included the use of the CCA as spatial filter, a layout optimization, that consisted in the cell arrangement of the virtual keyboard in a such way that neighbor cells were not much correlated in their VEP responses, and a stopping strategy, based on the difference of correlation between the most and second most correlated matching templates.

The proposed system achieved an average ITR of 48 bits/min in online experiments, showing the effectiveness of the reconvolution method combined to the optimized BCI setup.

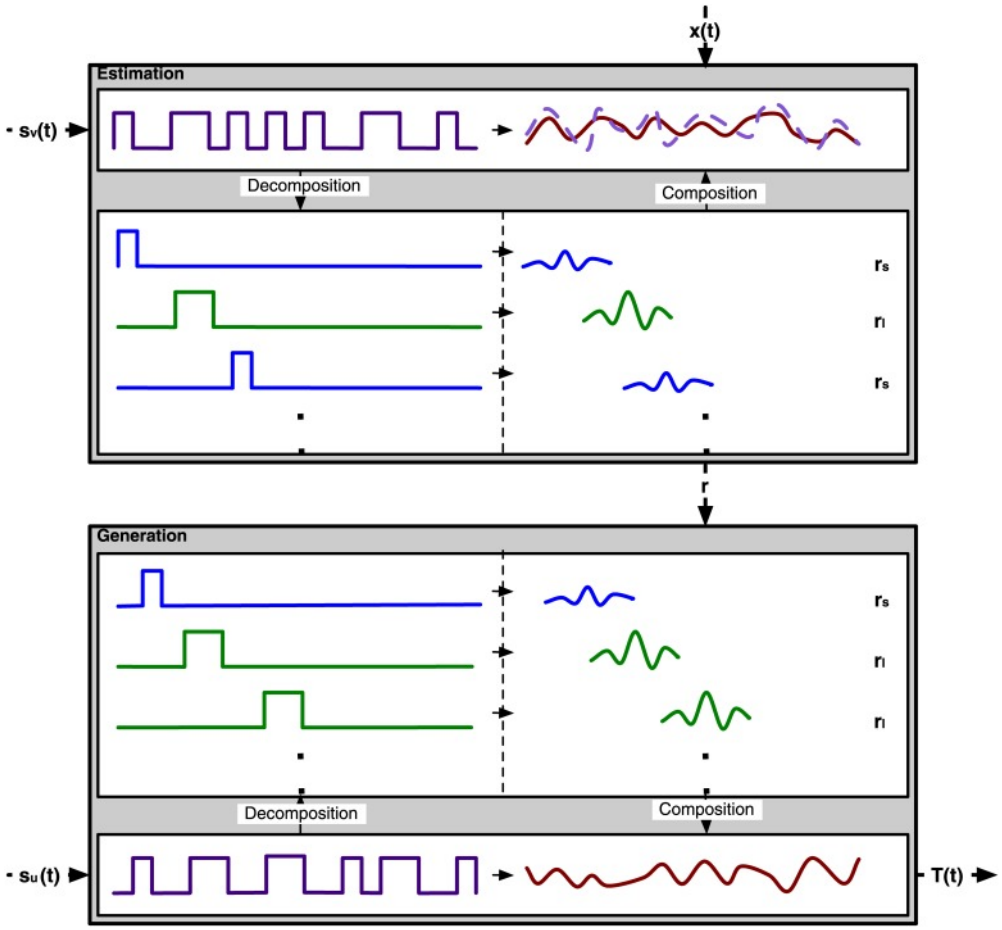


FIGURE 2.5 – Outline of the reconvolution method. Adapted from [Thielen et al., 2015].

Stimulus strategy

The stimulus strategy is crucial to build a high performance c-VEP BCI. Many studies investigate several parameters of stimulation to understand their influence in a c-VEP system, among them there are the stimulus color, stimulus presentation rate, number of stimulus cycles, stimulus length, type of stimulus sequence and stimulus proximity [Wei et al., 2016].

Many studies investigate the effect of stimulus specificity on the target flashing modulation, applying different pseudo-random sequences [Isaksen et al., 2017], with different bit length sequences [Wei et al., 2016]. [Isaksen et al., 2017] proved that among different types of sequence none provided a superior performance, showing that the "optimal-sequence" depends on the subject. [Wei et al., 2016] explored different stimuli layout parameters such as the size, color and proximity of the stimuli, different length sequences and different lags between adjacent stimuli, providing the best set of parameters to increase the performance in a multi-target c-VEP BCI. [Aminaka et al., 2015] propose a green-blue stimulus compared with the classical black-white, showing that the chromatic green–blue stimulus can give high result of accuracy, but not always better than the black-white color combination. [Nezamfar et al., 2015] explored different color stimulation: black and white, red and green and blue and yellow, stimulation sequence with three different bit lengths of 31, 63 and 127 bits but also three different bit presentation rates of 30, 60 and 110 bps. They showed that it is possible to find a compromise between high decision rate and subject comfort, using a m-sequence of 63-bit, a presentation rate of 60 bps and red-green for stimulus color.

About the stimulus presentation rate the most common value used for coding sequence is 60 Hz [Bin et al., 2011; Spüler et al., 2012]. [Wittevrongel et al., 2017] proposed a study in which they compare the traditional flashing pattern frequency of 60 Hz to a faster one at 120 Hz. Applying a novel decoding algorithm based on spatio-temporal beamforming, they showed that with a faster stimulation it is possible to increase the performance of the system, reaching an ITR of 172.87 bits/min.

However, analyzing these studies is clear that it is not possible to define an universal optimal stimulus parameter setting suitable for each BCI user in a c-VEP BCI system.

Type of keyboard

For a c-VEP speller two different keyboard layouts are used. The former is alphabetically virtual keyboard from A to Z followed by particulars character or/and numbers. Usually, there are 32 targets (8×4 grid) for a 63-bit stimulus sequence, the targets are separated by a black space and above the keyboard there is a text field showing the written text, as shown in Figure 2.2.

Another type of a virtual keyboard satisfies the principle of equivalent neighbors. The layout has 32 target (8×4 grid) surrounded by 28 non-targets, so that each target in the keyboard has eight neighboring targets. For each target, all neighbors keep fixed time lag relationships, as shown in Figure 2.6. This keyboard layout was designed by [Bin et al., 2011] following a previous study [Sutter, 1992] that demonstrated the influence on the VEP response of the stimuli outside the visual field. Therefore, the c-VEP combines evoked responses to the target as well as its neighbors. According to the principle of equivalent neighbors, c-VEPs corresponding to different targets are equivalent except for the time shift.

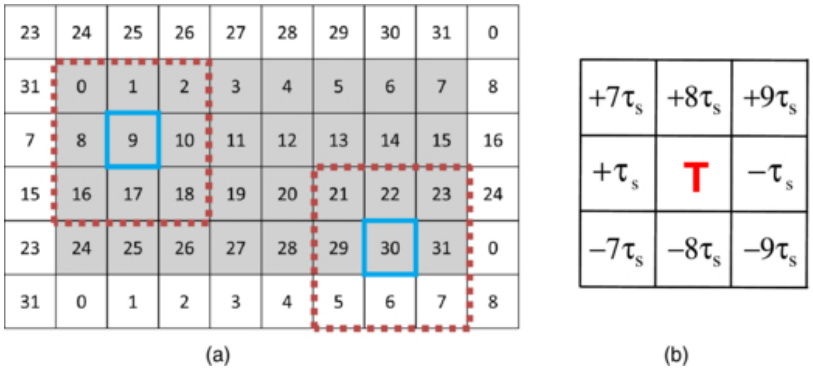


FIGURE 2.6 – The arrangement of the targets in the virtual keyboard following the principle of neighboring targets. Adapted from [Bin et al., 2011].

2.1.2 SSVEP speller

In a SSVEP BCI speller each target is flashed at different frequency eliciting evoked responses, that can be detected as a peak of band power in the EEG recorded over the occipital area (Figure 2.1) at the frequency of stimulation as well as its harmonics [Vialatte et al., 2010]. Hence, the target recognition is based on the detection of the responses at the same frequency.

Power Spectrum Density Analysis (PSDA) is a technique widely utilized for SSVEP detection, it consists in the extraction of the power spectrum and the frequencies of interest, using them as classifier features [Müller-Putz et al., 2005]. The PSDA computes the PSD values of EEG signals at different frequencies by spectral analysis techniques, such as Fast Fourier Transform (FFT) [Cheng et al., 2002], as shown in Figure 2.7. The frequency with the maximal PSD value is identified as the frequency of stimulation [Gao et al., 2019]. Then a classifier is applied to assign a frequency into a class. The most widely diffused classifier methods for SSVEP BCI are Support Vector Machine (SVM) and the Linear Discriminant Analysis (LDA) [Guger et al., 2012; Singla and Haseena, 2014; Oikonomou et al., 2016].

Recently, Canonical Correlation Analysis (CCA) was applied as a feature extraction method for SSVEP signal processing, enabling significant improvement of the performance, compared to PSDA [Wei et al., 2011]. The CCA was used to find optimal time-domain correlations between multi-channel EEG and reference sinusoidal signals [Bin et al., 2009b]. For SSVEP processing the CCA is utilized similarly to the c-VEP BCI processing, as explained in previous Paragraph 2.1.1. The difference is that c-VEP template signals are replaced with the sinusoidal reference signals Y_f at a frequency f, as shown in Figure 2.8. For each frequency of interest, the correlation between the EEG signal $\mathbf{X}$ and the reference signal Y_f is computed. This single correlation is used as a feature and the character that corresponds to the sinusoidal signal with the highest value of correlation with the EEG signal is identified as the target.

An important advantage of the CCA technique is that it avoids the calibration phase in SSVEP BCI systems. Indeed, it became a standard in calibration-free SSVEP-based BCIs [Bin et al., 2009b; Lin et al., 2006], thanks to the low impact of the noise and subject variability on the CCA technique. However, some works proposed the use of CCA with a calibration method [Zhang et al., 2014] in order to improve performance of the system or to develop an user-centred system.

Stimulus parameters

The stimulation frequency has a crucial influence on the amplitude of SSVEP response. For this reason it is fundamental to chose an adequate frequency

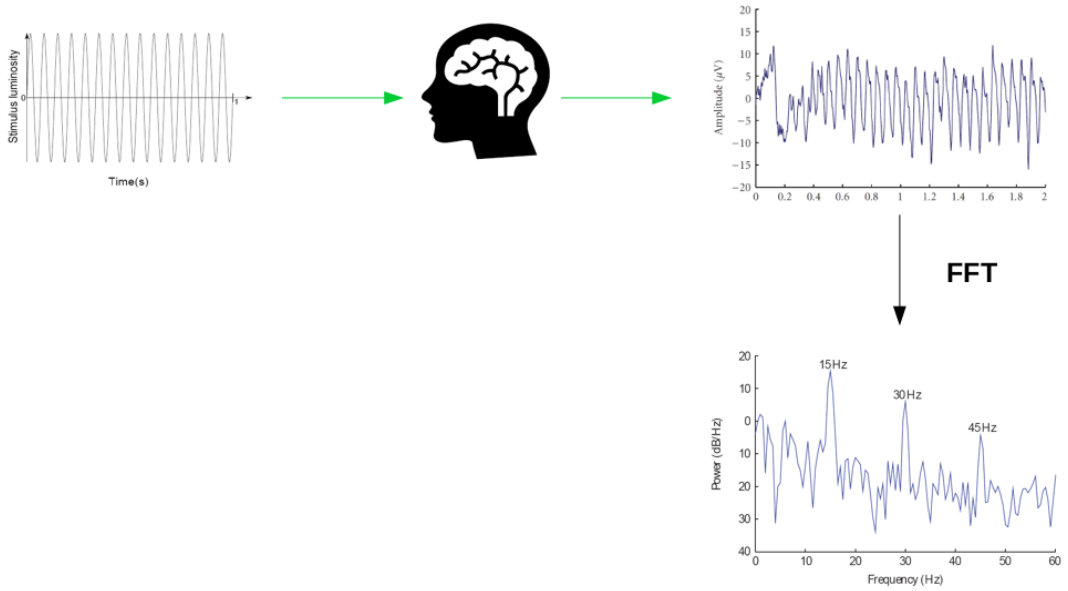


FIGURE 2.7 – An illustration of the SSVEP response. A stimulus at 15 Hz elicited a response on EEG of user, computing the FFT of that response yields a higher peak at 15 Hz and lower peaks at harmonics of the stimulus frequency. Adapted from [Materka and Poryzała, 2014].

of stimulation to design a performing SSVEP BCI. Nevertheless, the available stimulating frequencies are generally restricted by several factors.

First of all, not all stimulating frequencies always evoke high SSVEP responses [Jia et al., 2010] and the harmonics of some stimulation frequencies cannot be used simultaneously because they could interfere with one another, compromising the reliability of the evoked response [Volosyak et al., 2009]. The frequencies that can elicit strong SSVEP responses are highly dependent upon individuals as well as various environmental factors such as color, size, and contrast of the visual stimuli [Zhu et al., 2010; Jukiewicz and Cysewska-Sobusiak, 2016]. Many studies have investigated several ranges of frequency stimulation and it is reasonably accepted that a frequency of stimulation of 15 Hz evokes the highest SSVEP stimulation. Moreover, in the design of the GUI of a SSVEP speller the stimulation frequencies have to be set as divisors of the monitor refreshing rate to reach accurate SSVEP responses. For this reason, many existing SSVEP speller systems were implemented with only four or five stimulating frequencies and adopted "multi-step selection" strategies to spell each target [Cecotti, 2010; Volosyak, 2011]. In Figure 2.9 is shown an example of virtual keyboard developed for SSVEP

$$Y_f = \begin{pmatrix} \sin(2\pi f t) \\ \cos(2\pi f t) \\ \bullet \\ \bullet \\ \sin(2\pi N_h f t) \\ \cos(2\pi N_h f t) \end{pmatrix}$$

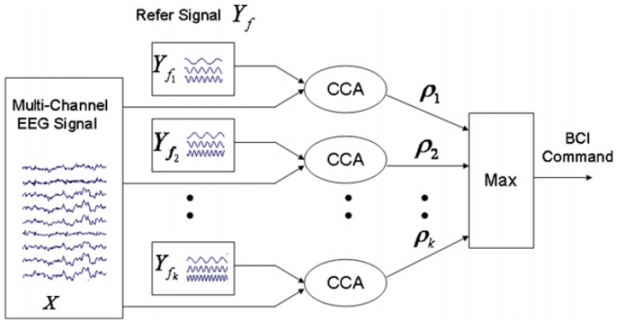


FIGURE 2.8 – Illustration of the application of CCA for SSVEP target detection. The CCA between a multi-channel EEG signal X and the reference signals Y_f for each SSVEP target is performed. The maximum correlation is selected as the classified target. Adapted from [Bin et al., 2009b].

speller following the "multistep selection" criteria.

Another important parameter that influences the performance of a SSVEP system is the color of the stimulation. Across the results of several studies it is not possible to define an universal stimulus color, because the level of comfort accompanying each color depends on the user's preference. Many works investigate the influence of the color white, green, red and blue in the SSVEP stimulation. Generally the white stimulation consistently leads to higher SSVEP amplitudes than colored stimulation. The reason of why white is better than the others color could be related to principle of the perception of light of the human eyes [Cao et al., 2012]. However, some works demonstrated that the white stimulation is not comfortable for the user, probably due to the high contrast with a black background [Bieger et al., 2010]. Among colored stimulations, red stimulation usually leads to higher amplitude than blue or green, but can be dangerous because can induct epileptic responses [Tello et al., 2015]. Thus, many works [Tello et al., 2015; Cao et al., 2012] proposed to chose the green color for a comfortable, safe and accurate SSVEP-BCI, followed by the blue color [Tello et al., 2015; Cao et al., 2012].

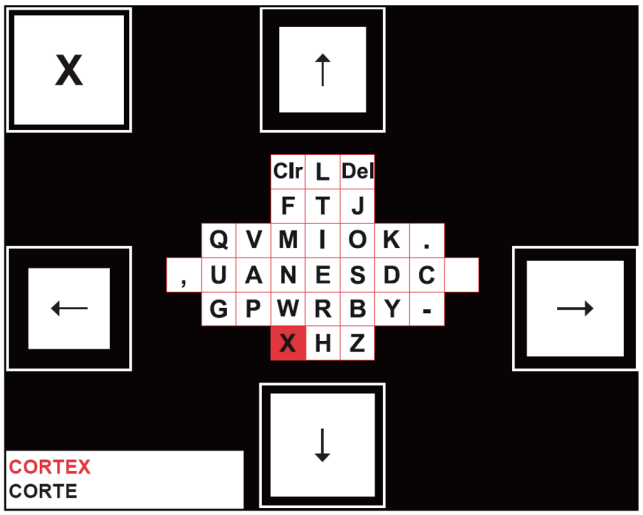


FIGURE 2.9 – Example of SSVEP virtual keyboard. The number of commands is limited to 5, the user can choose the character by moving the selection using the four arrows. The fifth command ("X") is for deleting a wrong selection. Adapted from [Volosyak, 2011].

2.1.3 P300 speller

The P300 speller is a device based on the detection of the P300 wave which is elicited by the expectation of a stimulus, as explained in Section 1.3. Groups of characters are flashed on the virtual keyboard and the user has to focus on the character that he/she wants to select, called *target character*. When the target character is flashed, the P300 wave occurs on the user's EEG and it means that the target character was part of the flashed group, as shown in Figure 2.10.

Typically the GUI of a P300 speller is a matrix of characters, all the characters are called dictionary D and a subset of characters of the dictionary is called sequence (S). The subset S_t is the group of characters that is flashed at time t. The P300 wave is detected analyzing the user's EEG signal and requires preprocessing, feature extraction and classification [Gayraud, 2018].

Usually, the EEG signal is preprocessed by a bandpass filtering with a low cutoff frequency of 0.1 Hz and high cutoff frequency lower than 30 Hz. Then the relevant information is extracted from the preprocessed EEG signal. The objective of feature extraction is to remove noise and other unnecessary information from the EEG signal in order to enhance the P300 response. To increase the signal-to-noise ratio some spatial filters techniques are proposed

in literature, such as discrete independent component analysis [Serby et al., 2005] and xDAWN spatial filter [Rivet et al., 2009]. Generally, the preprocessed EEG signal is divided into segments, starting from stimulus offset, with a specific number of time samples. These portions are called *epochs* and are constructed by concatenating samples over several channels. This representation defines the feature vector $\boldsymbol{x}_t$ for each stimulus presentation in time t.

Finally, the presence of the P300 response in a feature vector $\boldsymbol{x}_t$ can be estimated by a probabilistic classifier y_t. A probabilistic classifier returns values in the [0,1] interval and it estimates the probability of a data sample to belong to the category of *target epoch* or the other category *non-target epoch*. An example of probabilistic classifier largely diffused for P300 is the Linear Discriminant Analysis (LDA) [Hoffmann et al., 2008].

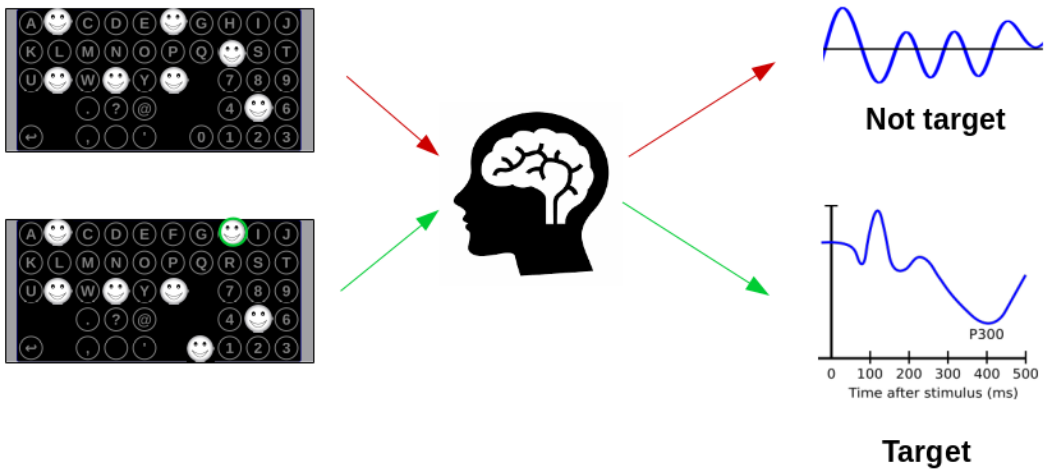


FIGURE 2.10 – P300 BCI. When the target character (character "H") is flashed, the P300 wave is recorded in the EEG of the subject.

Stimulation patterns

Many current researches on P300 speller interface concern the visual presentation of stimuli. The objective is to investigate the influence of several stimulus parameters to define a possible universal set of optimal parameters that increases the signal-to-noise ratio of the P300 response and consequently increasing the accuracy of the system. Many studies investigate the influence

of appearance of the virtual keyboard in terms of arrangements of the targets, such as the size of the matrix of characters. [Allison and Pineda, 2003] tested three different sizes of matrix (4×4, 8×8 and 12×12) showing that the P300 amplitude increased when the size of the matrix increased.

[Sellers et al., 2006] examined the effect of the matrix size combined to the duration of the Inter-Stimulus Interval (ISI) between two sequences of flashes. They crossed the matrix size of 3×3 and 6×6 with the ISI values of 175 ms and 350 ms, demonstrating that the accuracy reached with an ISI of 175 ms is higher than with 350 ms. Nevertheless, they showed that the matrix size do not have a high impact on the performance of the system. [Lu et al., 2013] proposed a study in which they analyzed the impact of stimulus presentation rate, such as the Inter-Stimulus Interval and the flash duration on the performance of the P300 speller. They found that the accuracy of the speller increased as the ISI increased and the flash duration has a minimal influence on the performance of the system. [Gonsalvez and Polich, 2002] used ISI values of 100, 200 and 400 ms and they confirmed that the P300 magnitude response increases with longer Inter-Stimulus Intervals. Nevertheless, they showed the importance of the Target-to-Target Interval (TTI).

The Target-to-Target Interval is an important parameter to avoid several issues that can influence the amplitude of the P300 response, such as the loss of attention of the user, that occurs if the intervals between two targets are less than 500 ms [Raymond et al., 1992], and the habituation of the user that is also influenced by the occurrence of the target flash. In fact the incorrect occurrence of the target flash can cause the problem of repetition blindness. This problem appears when two identical targets in a stream of non-targets are flashed at intervals between 100 to 500 ms, the second target may be missed [Kanwisher, 1987]. To avoid this problem [Townsend et al., 2010] proposed a checkerboard flashing paradigm (CBP) to replace the commonly used row-columns paradigm (RCP) [Donchin and Coles, 1988], in which the individual rows and columns alternatively flash on the screen. In the CBP paradigm six different characters randomly located in the matrix flashed simultaneously instead of grouping characters by row or column, so the occurrence of characters that flash consecutively is reduced and adjacency errors caused by neighbors flashing consecutively are eliminated. The CBP method improved significantly the performance of the P300 speller reaching a ITR of 23 bits/min with respect to 19 bits/min reached with the RCP paradigm.

[Cecotti and Rivet, 2010] proposed a simple solution which minimizes adjacent distractors, randomizing the appearance of rows and columns flashed in the matrix. This results in groups which appear random on screen, but retain the properties of rows and columns. [Perrin et al., 2012], following the idea of randomization, created dynamic random stimulation (DRS) sequences, so after each stimulus cycle the flashing characters form more random groups than the original row-column sequences, but the group remain fixed across stimulus cycles.

[Thomas et al., 2014] proposed a novel strategy to generating optimal flashing sequences, called RIPRAND. The proposed strategy combines the idea of randomization of DRS sequences with an algorithm inspired by the restricted isometry property (RIP), generating sequences that share as few elements as possible among each other across a stimulus cycle, as shown in Figure 2.11. In this way the sequences of stimulation appear random to the user, preserving uniform flashing of the elements on the virtual keyboard, a reduction of double flashes and minimizing the elements that flashed together repeatedly. Moreover, they improved RIPRAND through of an algorithm of early stopping, increasing the discriminability between the different characters accumulating evidence for individual characters. The results reached showed that the RIPRAND algorithm with early stopping increased the accuracy of the P300 speller with respect to other flashing strategies, such as the row-column paradigm [Thomas et al., 2013].

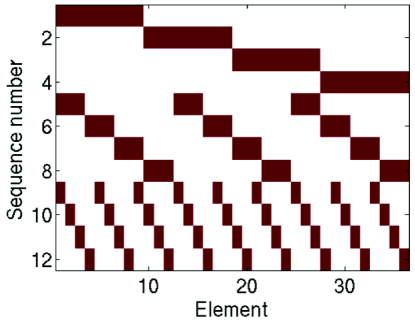


FIGURE 2.11 – Example of flashing sequences (not randomized) generated for one repetition with a flash ratio rate of 1/4 for a 36 element grid. Red indicates flashed elements. Adapted from [Thomas et al., 2013].

2.1.4 Limitations and research

There are many limitations that prevent the diffusion of the EEG-BCI spellers in real applications. These limitations can be considered as challenges for the BCI research to find solutions to improve the diffusion of the system outside the laboratory. In particular, it is fundamental to find solutions that allow impaired people to use every day a BCI system to really improve the quality of their daily-life. Among these limitations there is the tedious calibration phase, that is a consequence of the problem of the inter-subject and inter-session variability.

Many solutions and investigations are proposed to tackle these limitations, such as the improvement of the classification algorithm with machine learning methods or the incorporation of a language model.

Table 2.2 gives a comparison of the main features and limitations that characterize the spellers presented in Section 2.1. In the following sections the calibration problem will be discussed with a focus on the solutions proposed in the literature.

Calibration

Both c-VEP BCI speller and P300 speller need a calibration phase, in which the system sets the preprocessing and calibration parameters to extract the relevant information from the EEG signal, in term of spatial filters and classifiers. This phase is fundamental but is also really annoying, long and tedious for the subject. It is considered time consuming and one of the limitations of the diffusion of the BCI system in real applications [Millán et al., 2010]. A solution to this limitation is the development of a system that avoids the calibration phase. Many studies underline the importance of a calibration-free BCI and propose different strategies to reduce the calibration time duration.

Many works [Blankertz et al., 2010] proposed an approach based on transfer learning methods [Jayaram et al., 2016], in which the classifier is pre-trained with previous available data of the same or different subjects.

[Lu et al., 2009] proposed an adaptive online learning method to a P300 speller. The first step of the method consists in defining offline a generic subject-independent model from a EEG data set of several subjects, in order

to find the specific structure of the P300 response. Then, with a new subject the system automatically captures subject-specific EEG characteristics during online operation without a supervised calibration. The final decision is done by labels predicted by either the subject-independent model or the adapted subject-specific model, depending on a confidence score. They demonstrated that after the spelling of 10–20 characters with the online adaptation, the accuracy of the adapted model reached the value of a trained supervised subject-specific model. To assess the strategy proposed by [Lu et al., 2009], [Jin et al., 2013] proposed an experimental protocol in which five different classification models were tested online, including a generic model and a generic model combined with the online training strategy, in order to find the optimal training strategy. The reached results showed that the averaged accuracy increased with a calibration that combines the generic model and the online training with respect to the generic model used alone.

[Kindermans et al., 2012b] proposed an unsupervised model incorporating a transfer learning method and a language model to a P300 speller BCI. They demonstrated that this unsupervised model can reach even better performance than the supervised model.

Recently unsupervised calibration methods have started spreading in the BCI community. In this approach the classifier is pre-trained with previous available data of the same or different subjects. The latter approach is related to the unsupervised learning in which the classifier has no prior knowledge from other training subjects, but it has to learn the relevant brain features from scratch. [Kindermans et al., 2012a]introduced the first unsupervised classification method for a P300 system, that proposed an approach based on the Expectation Maximization (EM) algorithm in order to find the parameter of a linear classifier.

[Hübner et al., 2017] proposed a modification of the P300 spelling interface to apply on a BCI the learning from label proportions (LLP) [Quadrianto et al., 2009]. Then, [Verhoeven et al., 2017] combined LLP and EM strategies, and this mixed method was verified by [Huebner et al., 2018] in an online study showing that the combination of the two strategies performs better than the unsupervised classifier applied alone, but the limit of the method is that it is applicable only to Evoked Related Potential (ERP) data, suggesting that the future research in this topic should take in account also the user and the

interface.

For an extensive reviews on unsupervised adaptation and unsupervised classification we suggest to refer to [Huebner et al., 2018].

Language model

Many BCI systems exploit a language model to improve the performance of the systems [Mora-Cortes et al., 2014]. The incorporation of a language model into a BCI speller can impact various aspects of the system, from changes in the GUI to modifying the signal classification algorithms.

The GUI of a BCI speller system can be improved by incorporating language information from a word prediction model. A word prediction system uses dictionaries and after each character selection, the system looks for the partially completed word and returns potential target words. The most frequent words among them are presented to a target list from which the user can select the possible target word. The user has the option of continuing to spell by selecting individual characters, or selecting one of the proposed words, so that the user is able to spell faster, as the word does not always need to be typed character-by-character. Many works design GUI incorporating the word prediction model and integrating the word suggestion list that can be either on a separate layout [Ryan et al., 2010] than the one for character-by-character spelling, or in the ordinary layout [Kaufmann et al., 2012]. An example of the two different layouts is shown in Figure 2.13. The integration of the language model in the ordinary layout can be more practical for the user because it does not require an additional BCI command to switch between two layouts. In summary, word prediction allows to increase typing speed by reducing the number of selections required.

Like word prediction, also incorporating a language model into the classifier can improve the performance of the spelling system, increasing the accuracy and reducing the amount of time required for single selections. Classifiers based on language model use prior characters to generate a prior probability distribution for the next character to be selected, based on the previously typed character or text. Usually, this prior probability distribution is constructed from a dictionary or a corpus of text, the suggested sequence of characters or words are updated after each character selection following an

n-gram prediction model. N-grams are simply all combinations of adjacent words or letters of length n that can be found in a source text. This model considers combinations of adjacent letters or words of length n from a dictionary or a text, each item x_i has the probability $P(x_t|x_{t-1}...x_0)$ and each letter or word is selected following the previously typed character or word. For example, given a character, c, it possible to construct a list of n-grams from c by finding pairs of characters that occur next to each other. Several types of n-grams are incorporate and tested in BCI spellers [Orhan et al., 2012]. Many works [Samizo et al., 2013; Speier et al., 2011] reported an improvement of the speed and accuracy of typing with the P300 speller integrated with n-grams language. [Gembler and Volosyak, 2019] developed a dictionary-driven c-VEP BCI, proving that the incorporation of bi-gram word prediction allows to increase the performance of a c-VEP system. The GUI designed by [Gembler and Volosyak, 2019] is illustrated in Figure 2.12.

Moreover, thanks to the prior language information it is possible to include other features in the system, such as word completion, automatic error correction or dynamic stopping [Speier et al., 2016]. Finally, all these features applied alone or combined can effectively improve accuracy and typing speed of BCI spellers [Speier et al., 2016].

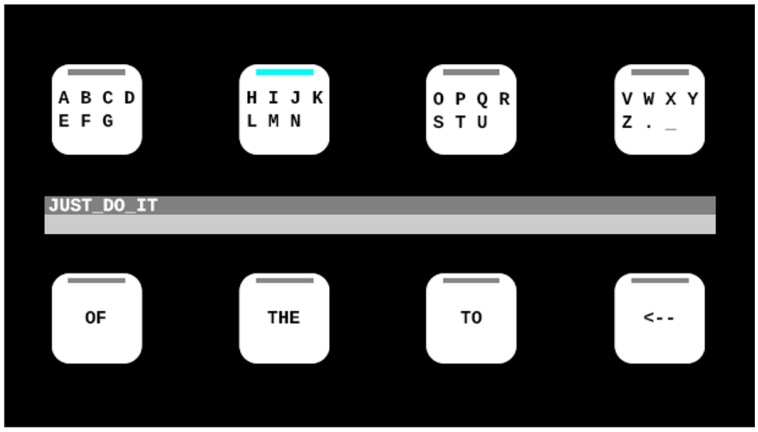


FIGURE 2.12 – C-VEP speller with incorporation of a language model in the classification algorithm. In the upper part of the screen the user can select the group in which there is the desired character. In the lower part of the screen the list of the words selected by the n-gram language model integrated in the classifier are proposed to the user. The user can directly select among the suggested words, if the target one is among them, or can continue the spelling letter-by-letter. Adapted from [Gembler and Volosyak, 2019].

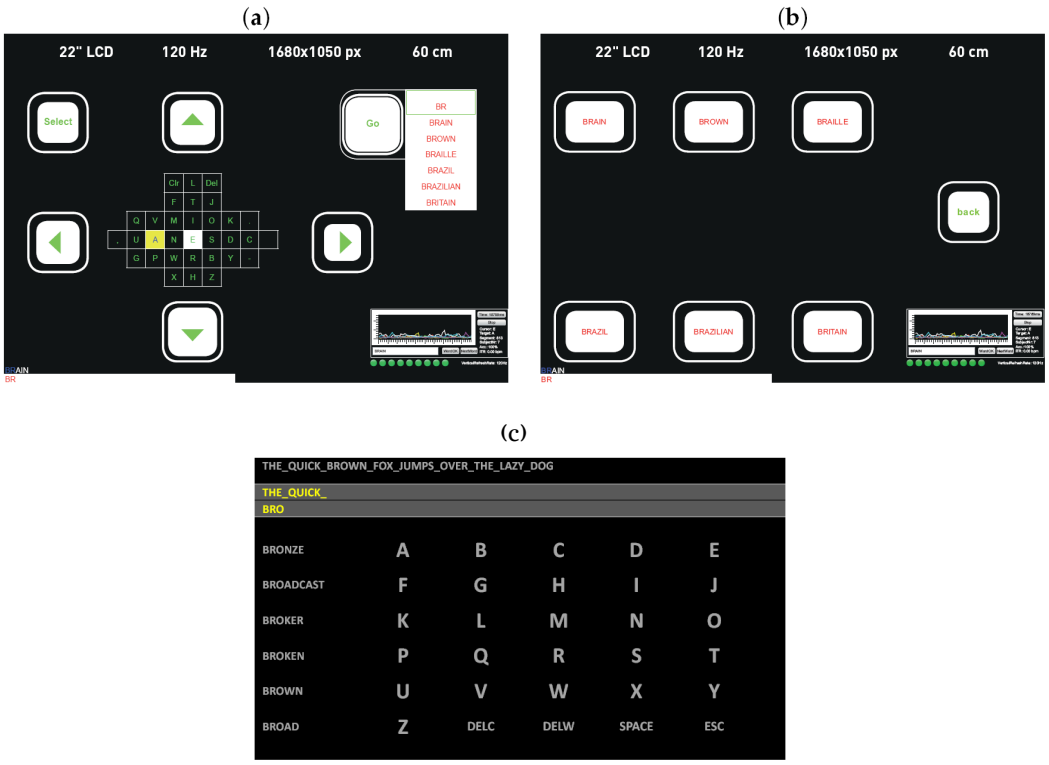


FIGURE 2.13 – Example of GUI incorporating the a language model. (a) The original SSVEP speller (shown in Figure 2.7) is modified by incorporating a dictionary in the GUI (b) The second stage of the GUI, where the list of suggested words are presented to the user, who can choose the target word. (c) GUI of a P300 speller incorporating the dictionary for word suggestion in the interface. Figure adapted from [Rezeika et al., 2018]

2.2 Mental Imagery based BCI

In a BCI based on Mental Imagery (MI) the user has to perform a specific mental task that generates EEG components that can be translated in commands for a BCI system. There are many applications that involve MI BCI, such as neurorehabilitation [Van Dokkum et al., 2015], control of external devices [Cincotti et al., 2008], virtual reality [Leeb et al., 2007] and gaming [Kauhanen et al., 2007].

Among mental tasks the cognitive process of imagining moving a part of the body produces the synchronization on the neurons of motor area (shown in Figure 2.14), as explained in Section 1.3. Motor imagination is the most widely applied mental task in the BCI field. The user can imagine to move a

part of the body, such as a foot or a hand, based on visual cues. The mental tasks include also other cognitive tasks, such as:

— Visual counting: where the subjects mentally count incrementally, starting from a number.
— Mental subtraction: where the subjects repeat mentally subtraction between numbers (e.g. 257–9 = ..., 248–9 = ..., 239–9 = ...).
— Mental rotation: where the subject has to imagine rotation of a three dimensional geometric object around an axis.
— Lexical imagination: generating as many words as possible beginning with a specific letter.
— Auditory imagination: the subject has to imagine to sing a song.

For sensorimotor BCI the feature extraction and classification step are crucial to find discriminative and informative features to translate in commands that carry out user's intent. The literature proposes different methods applied in the features extraction phase. The most used are the bandpower extraction and the Common Spatial Pattern (CSP). For the classification phase many classification algorithms are devised such as support vector machine (SVM) and linear discriminate analysis (LDA). For an extensive review of dominant signal processing algorithms in sensorimotor BCIs refer to [Aggarwal and Chugh, 2019]

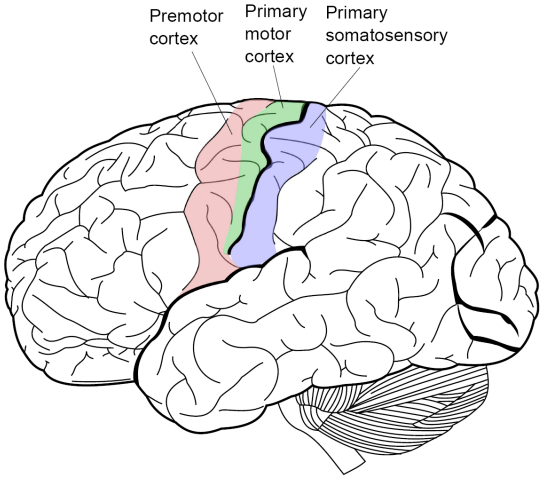


FIGURE 2.14 – Motor and sensorimotor cortical areas. Adapted from [Evain, 2016].

2.2.1 Limitations and research

One of the most important limitations for SMR-based BCIs is the high intra- and inter-subject variability, preventing the use of such systems in common life [Saha and Baumert, 2019]. The experimental setting, psychological and neurophysiological parameters influence the SMR that often varies over time and across subjects, since the motor learning process differs across healthy subject and patients, but also among patients with motor impairments. This can be the consequence of different causes that impact inter-subject variability for cognitive and neurological factors. Emotional factors such as fatigue, memory load, attention and reaction time modulate instantaneous brain activity, and can affect performance of SMR-BCI systems.

Another important parameter impacted by this variability is the design of a SMR system for an impaired subject, it has to take into account a phase of definition of the tasks most adapted to the subject, considering his/her neurological response but also his/her possibility to carry out specific tasks.

It is widely agreed that in order to use SMR-based BCI systems training is fundamental, moreover to further improve the skill of modulating sensorimotor rhythms [Wolpaw and Wolpaw, 2012] an even more substantial training is required. Nevertheless, the basic mechanism of SMR learning is not clear. Many studies investigated on the motor learning process that promote plasticity in the sensorimotor networks and improve both motor and perceptual skills [Ostry and Gribble, 2016] proving that BCI skill acquisition effectively allows to improve the BCI performance also in impaired subjects. Yet, subject-specific training sessions may be required because the induction of plasticity varies significantly across subjects [Saha and Baumert, 2019].

2.3 Conclusions

This chapter presented some applications of BCI, analyzing the limitations and the possible solutions. In Table 2.1 the main limitations discussed in this chapter are summarized.

Some BCI speller applications are presented describing their pipeline and referring to the state-of-the-art, in order to outline the methods generally used

in existing implementations. The main features of the presented BCI spellers are summarized in Table 2.2. A particular focus was put to highlight the importance of the GUI in a speller, analyzing key parameters of each stimulation paradigm. Something that all such paradigms share is that there is no set of parameters universally valid for any of them, because there are many factors influencing the cerebral response, the most important one being the condition of the user. In particular, a focus of this manuscript will be on the investigation of the influence of the parameters of stimulation for a c-VEP BCI system implemented from scratch, in order to effectively evaluate the influence of the stimulation strategy, that could impact the intra/inter-subject variability.

Moreover we have discussed the limitation caused by the calibration phase, which is still necessary for the proper operation of many BCI systems. We have also carried out an extensive analysis of methods proposed in literature to avoid this expensive calibration phase.

Finally, the MI-based BCI system is presented in this chapter. Such type of system presents the same limitations as the ones presented earlier, that is the importance of creating an user-centered system in order to obtain an increase of the performance of the system.

In forthcoming sections, we will present the methods to tackle some limitations of BCI systems, with the goal of stimulating the development of user-centered systems.

TABLE 2.1 – Summary of the main advantages and limitations of the BCI applications presented in Chapter 2.

	ERP and VEP BCI spellers	**MI-BCI**
Limitations	Time-consuming calibration phase High influence of stimulation strategy Permanent attention to external stimuli Influence of external factors Inter/intra-subject variability	Very time-consuming training (months or weeks) Not all users are able to obtain control Need to define specific tasks for each user Influence of external factors Inter/intra-subject variability

TABLE 2.2 – Summary of the main features of the BCI spellers presented in Section 2.1

BCI paradigm	**c-VEP**	**SSVEP**	**P300**
Brain activity pattern	Electrical potential in the visual cortex	Electrical potential in the visual cortex	Positive deflection in the parietal region 300 ms after the infrequent attended stimulus
Type of stimulation	All characters flash simultaneously according to a binary sequence	All characters flash simultaneously at different frequencies	Group/sequence of characters flash together
Relevant stimulus parameters	Type of sequence Stimulation color Stimulation frequency Keyboard arrangement	Stimulation frequency Stimulation color Keyboard arrangement	ISI TTI Sequence generation Keyboard arrangement
Limitations	Intra/inter-subject variability Calibration phase required High influence of stimulation strategy	Not suitable for applications with many options High influence of stimulation strategy	Intra/inter-subject variability Calibration phase required High influence of stimulation strategy
ITR	60–140 bits/min	30–100 bits/min	20–35 bits/min

Part III

Contributed Methods

3 Design and Flashing Strategy for BCI Spellers

In the previous chapter we reviewed the VEP BCI spellers. In particular, we underlined the importance of the speller GUI and the lack of an universal set of stimulus parameters.

In this chapter we will present two pilot studies to investigate and implement different stimulus strategies in BCI spellers. These studies show the impact of some stimulus parameters on the performance of a BCI speller, ultimately allowing to design a system with higher performance.

The first work concerns the implementation of a flashing strategy in a P300 system in order to improve the discriminability between flashed characters and their influence on the performance of a P300 speller. The method, the experimental setup and procedure will be described in the next section, followed by the results and deduced conclusions.

The results obtained with the implementation of this flashing strategy led our work to move to the design and development a GUI for a c-VEP BCI speller, that will be the subject of the second work presented in this chapter. It concerns the study and development of a GUI for a c-VEP BCI speller, aiming to provide some new features for a BCI speller GUI and software starting from scratch. Finally, different experimental setups are tested and compared to prove the effectiveness of the developed system.

3.1 Flashing strategy for a P300 speller

As discussed in Section 2.1.3, flashing strategies are fundamental to increase the performance of a P300 speller. The work detailed in this section aims at investigating the effect of increasing the discriminability between different

characters in the P300 paradigm. The BCI system developed by [Thomas et al., 2014] was the starting point of this work. The objective of the proposed flashing method is to increase the discriminability and detection speed of the target character in a P300 speller, using a letter-based flashing strategy instead of a group-based flashing strategy.

3.1.1 Letter-based flashing strategy

The letter-based flashing strategy is an enhancement of the group-based flashing strategy proposed by [Thomas et al., 2014] in order to improve the discriminability of the target character.

The stimulation strategy proposed by [Thomas et al., 2014] is a group-based flashing strategy, in which all characters k of each flashing sequence S_i ($i = 1, ..., n$ where n is the number of sequences per repetition [Thomas et al., 2014]) are flashed together, as shown in Figure 3.1 and explained in Section 2.1.3.

The proposed letter-based flashing strategy employs as sequence of stimulation the sequence of characters generated by the RIPRAND method, in order to preserve its benefits, but it modifies the flashing strategy proposed by [Thomas et al., 2014] during the calibration phase, changing the flashing presentation and tag synchronization.

Flashing presentation

In the letter-based flashing strategy the characters k of a flashing sequence S_i are flashed one after the other, as shown in Figure 3.2, and not all together as in the group-based flashing strategy proposed by [Thomas et al., 2014]. Each character k belonging to a sequence S_i is flashed for a time duration $t = t_{character} + t_{overlap}$, where $t_{character}$ corresponds to two frames (32 ms for a screen with refresh rate of 60 Hz) and it is the flashing duration in which the character is flashed alone on the virtual keyboard, while $t_{overlap}$ corresponds to one frame (16 ms for a screen with refresh rate of 60 Hz). During the interval $t_{overlap}$ the first frame of a new character and the previous character are both flashed in order to increase the length of the flashing of every character

k. As a consequence the two characters are overlapped for a frame, as we can see in Figure 3.2.

Tag synchronization

The letter-based flashing strategy was integrated in the calibration phase, as mentioned before. In this phase a trigger signal is provided by a TCP network connection to synchronize the stimulus presentation and the EEG data recordings. The proposed method allows to improve the discriminability by sending a tag to the signal acquisition module of the system in correspondence of the first flashing frame of each character k, along with a label that indicates if the character is a target or not, as shown in Figure 3.2. Therefore, the definition of target and non-target characters is immediately available at the preprocessing phase, in which the signal is divided in epochs (as explained in Section 2.1.3), since the instant in which the target character flashes is known.

On the contrary, during the calibration phase of the group-flashed strategy a tag is sent at the beginning of each sequence S_i to indicate if the target character is present in the sequence or not, as shown in Figure 3.1, therefore the whole group flashing with the target character is labeled target.

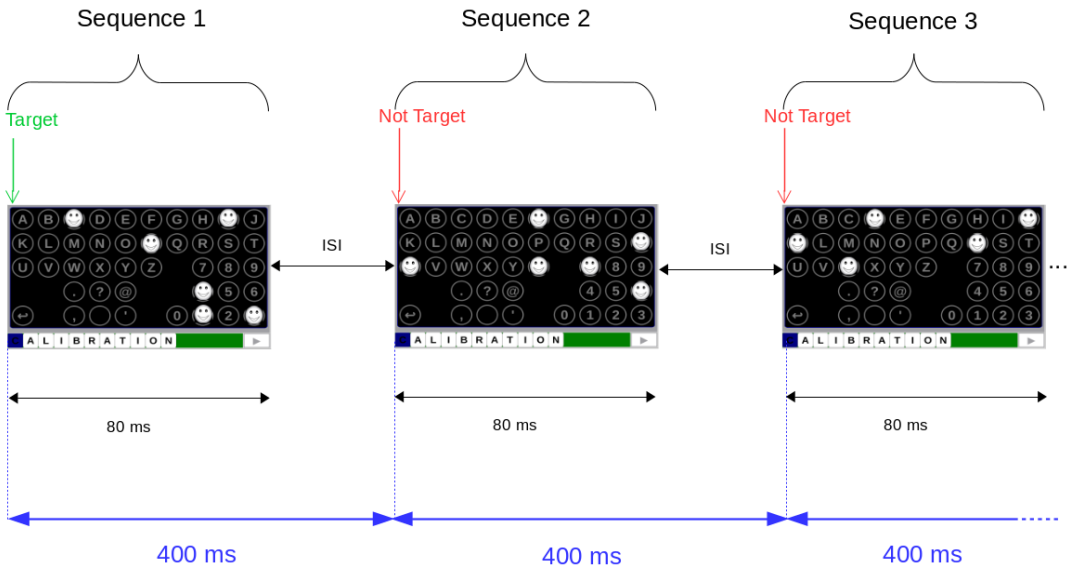


FIGURE 3.1 – Example of group-based flashing method. The first three sequences of characters, of one repetition each, are flashed for 400 ms (ISI interval included). A label is sent to the acquisition system at the beginning of each stimulus to indicate if the target character is present or not in the flashed group.

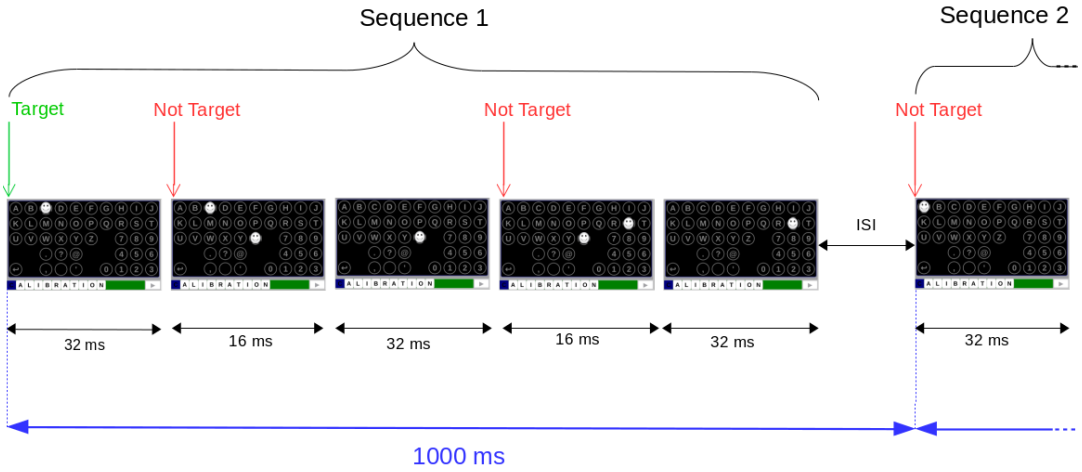


FIGURE 3.2 – Example of letter-based flashing method. The sequence S_1 = "C Z S" is flashed and the character "C" is the target. The first and the last letter are flashed for 3 frames and the other ones for 4 frames in total (1 frame = 16 ms with a screen at 60 Hz). A single sequence is flashed for a total duration of 1000 ms (ISI interval included). Then, the first character of the next sequence will be flashed. At each new character a label is sent to the acquisition system, to indicate if the character is the target (green arrow) or not (red arrow).

3.1.2 Experimental data

Two healthy volunteers (one male of 28 years old and a female of 26 years old) took part at the P300 experiment. All subjects had normal or correct to normal vision and did not suffer from epilepsy or other nervous diseases. Both subjects completed the whole experiment. Only one subject was P300 BCI-naive participant. During each session the EEG signal of the subject was recorded from a ANT-Waveguard cap and a Refa8 amplifier (512 Hz sampling rate). To set the impedance between the electrodes and the subject's skin below 10 kΩ, a conductive gel was applied to the ground (FPz) and to the 12 electrodes placed in positions Fz, F3, F4, Cz, C3, C4, Pz, P7, P8, Oz, O1, O2.

The P300 speller system was designed by [Thomas et al., 2014]. It consisted of OpenViBE [Renard et al., 2010] for signal acquisition and a custom keyboard-display control software, both running on a Windows 7 laptop. The laptop screen was used to monitor the EEG signal quality during acquisition, while the virtual keyboard was displayed on a separate LCD monitor. The virtual keyboard layout (Figure 3.2) consisted in a 10×5 matrix with 43 characters: letters sorted alphabetically from A to Z followed by backspace, symbols " . ", " ? ", " @ ", " , ", " ' ", UNDO and numbers from 0 to 9.

Each character was placed in a circle with a black background. Below the matrix of characters a text field showed the characters of the word that the subject had to gaze at. The flashing consisted of briefly covering the character on the virtual keyboard with a "smiley face", which elicits stronger P300 responses than simply highlighting the character [Jin et al., 2012; Guy et al., 2018]. A trigger signal was provided by a TCP network connection to synchronize the stimulus presentation and the EEG data recordings.

The stimulation sequences S_i were generated following the RIPRAND method [Thomas et al., 2014] with a flash-rate of 1/4 and a number of repetitions equal to 3, flashed following the letter-based method. The interface was used in calibration mode and the word to spell was "CALIBRATION". Moreover during the experimental session the subjects had to count in their mind how many times the target letter was flashed, to keep their attention focused on the experiment.

Offline processing

The recorded EEG signal was filtered between 1 Hz and 20 Hz with a 5th order Butterworth bandpass filter, in order to remove as much as possible the external interference and possible artifacts. Then the signal was divided in epochs of 500 ms following the tag sent during the experiment.

The XDAWN spatial filter [Rivet et al., 2009] was applied to reduce the number of channels from 12 to 3, allowing to obtain a linear combination of channels increasing the signal-to-noise ratio. As last step, a binary classification was applied by a LDA classifier [Hoffmann et al., 2008]. In order to carry out the classification it is necessary to divide the dataset of each subject in a *training set* and a *test set*, therefore the first 70 % of epochs of signal were considered as the *training set* and the remaining 30 % as the *test set*.

The LDA classifier was applied considering two different sample sizes. The former, called LDA_{all}, to classify all the sequences of characters S_i flashed during each experimental session, hence including flashing sequences with and without target characters. The latter, called LDA_{target} to classify only the flashing sequences of characters S_i which included the target character k^*. Finally, a ROC curve was applied to assess the performance of the classification.

ROC curve

The Receiver Operating Characteristic (ROC) is a method that returns a curve which is used to evaluate the performance of a binary classification, that can return four outcomes: True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) within a confusion matrix [Hamadicharef, 2010]. The ROC curve is created by plotting true positives of the classification with respect to false positives, corresponding to different threshold values. The TPR and and TNR are the performance measures of the classifier and are defined as follows:

— The True Positive Rate (TPR), or sensitivity, is the proportion of correctly identified positive outcomes of the classifier $TPR = \dfrac{TP}{TP + FN}$, where TP indicates true positive and FN false negative.

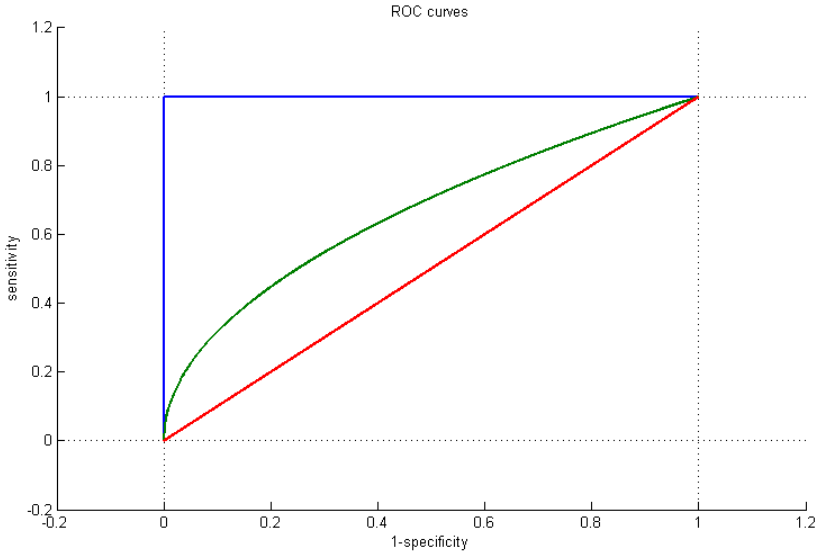


FIGURE 3.3 – Example of ROC curves: perfect test (blue), non significant test (red) and a realistic, not very accurate, test (green).

— The True Negative Rate (TNR), or specificity, is the proportion of correctly identified negative outcomes of the classifier $TNR = \frac{TN}{FP + TN}$, where TN indicates true negative and FP false positive.

The integral of the ROC curve, called AUC (Area Under the Curve), measures the discrimination, that is the ability of the test to correctly classify each sample between two classes. Such an area has values limited between 0.5 and 1, the higher this value, the better is the discrimination of the test. Practically, more the ROC approaches an ideal step curve the better the discrimination, as shown in Figure 3.3.

In order to understand the AUC values we refer to the classification proposed by [Swets, 1988]: if AUC $= 0.5$ the test is not significant, if $0.5 <$ AUC ≤ 0.7 the test is not very accurate, if $0.7 <$ AUC ≤ 0.9 the test is moderately accurate, if $0.9 <$ AUC < 1 the test is very accurate and the test is perfect only if AUC $= 1$.

3.1.3 Results

Figure 3.4 shows the average of the target epochs and non-target epochs computed over the three occipital channels (O1, Oz, O2), for which the P300 response is the most prominent. The curves show that for both subjects we obtained a response coherent with respect to the division in target and not-target epochs, following the tag sent during the acquisition. Both subjects show a pronounced P300 waveform for target epochs and the absence of it for not-target epochs, proving the good discriminability achieved by a letter-based flashing strategy.

Moreover, to evaluate the performance of the deployed binary classification methods, ROC curve for both subjects are plotted and shown in Figure 3.5. We can notice that the area under the curve of the subject S1 is larger with the LDA_{all} method than with the LDA_{target} method. This is also found from the values of the AUC listed in Table 3.1, that are respectively 96 % and 94 %.

Similarly, the difference among the ROC curves of the two methods is small also for the subject S2, as confirmed by the values of AUC reported in Table 3.1, respectively 87 % with LDA_{target} and 88 % wih LDA_{all}.

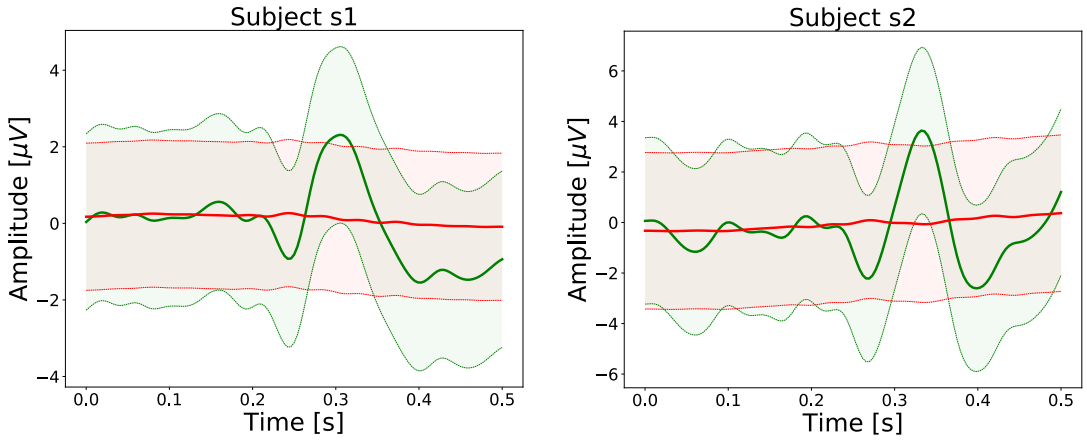


FIGURE 3.4 – Mean non-target epoch (red) and mean target epoch (green) and standard deviations, which shows the P300 response around 0.3 s.

TABLE 3.1 – AUC with LDA_{target} and LDA_{all} methods.

Method	Subject	#target epochs	#all epochs	AUC
LDA_{target}	S1	94	1015	94%
	S2	91	989	87%
LDA_{all}	S1	94	3862	96%
	S2	91	3862	88%

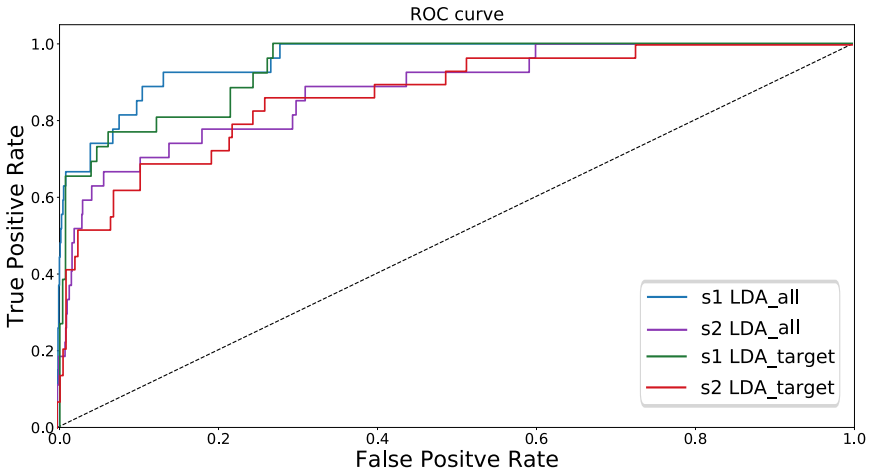


FIGURE 3.5 – ROC curves calculated considering all character sequences, with and without the target characters (blue for subject S1 and purple for subject S2), and considering only the sequences with the target character (green for subject S1 and red for subject S2). The black dash line indicates the 50% chance level for the binary classifier.

3.1.4 Discussion and perspectives

Beyond the results shown in Figure 3.5 and detailed in Table 3.1, we can observe the importance of the experience of the user. Indeed the best results were reached for subject S1, who already took part in many experiments with the P300 speller. This could mean that to obtain a good performance the subject involvement is fundamental.

For both subjects the results reached are not different between the two methods. It means that it is possible to define the P300 response during the calibration phase with a smaller sample size, that includes only the flashing sequences of characters with the target epochs. Indeed in Table 3.1 we can see that with the LDA_{target} method the total number of samples is 1015 for subject S1 and 989 for subject S2. On the other hand with the LDA_{all} method the number of samples is 3862 for both subjects. The increase of discriminability with the letter-based flashing method has proven that it is possible to obtain good classification results also with a smaller amount of stimulation sequences.

Even if the preliminary results open the possibility for the development of a more precise flashing strategy for the detection of the target response, the issue is that the system becomes slower, as shown in Table 3.2. This is due to the fact that, to increase the discriminability, each target is flashed individually, as shown in Figure 3.2. Therefore if we compare the letter-based flashing strategy with respect to the group-based flashing strategy (Figure 3.1) the time required to flash a stimulation sequence is 128 ms and 80 ms, respectively. Moreover the total flashing duration of a sequence of stimulation is longer with the letter-based flashing strategy: 1000 ms against the 400 ms of the group-based flashing strategy. In the letter-based strategy the flashing duration increased in order to ensure the correct tag synchronization with the system, indeed we tried to decrease the total flashing duration but problems with tag synchronization occurred.

Crucially, a longer stimulation implies a lower communication performance in the actual use of the system, caused by the experimental protocol becoming longer and uncomfortable for the subject. For this reason we changed the focus of our research moving to a different type of stimulus strategy which aims at achieving high target discriminability but also high speed at the same time. This strategy is the c-VEP BCI.

Indeed, comparing the c-VEP stimulus strategy with the letter-based flashing strategy, the discriminability level will be high for both methods. In the former method there is a specific template for each character, while in the latter method we can know the specific epoch of the target character. Moreover, in Table 3.2 we report an example of the calibration time required to

the stimulus strategies analyzed in this section and we can notice that the average calibration time of the c-VEP stimulus strategy is shorter than the one needed by the letter-based flashing strategy.

TABLE 3.2 – Main metrics of the described stimulus strategies for VEP BCI. For both P300 strategies the calibration time is computed simulating the spelling of the word "CALIBRATION" with a flash-rate equal to 1/4 and 3 repetitions. While, for the c-VEP strategy we refer to the calibration phase reported by [Bin et al., 2011], in which the user gazed at a reference character for 200 repetitions.

	P300 group-based	**P300 letter-based**	**c-VEP**
Stimulus type	sequence of characters	sequence of characters	m-sequence
Calibration time	$\approx$ 3.5 min	$\approx$ 9 min	$\approx$ 3.5 min
Discriminability	low	high	high

3.2 Design of a c-VEP BCI

In this section a preliminary design of a c-VEP BCI will be presented. The designed system is integrated in the P300 speller software developed by [Thomas et al., 2014], hence the user can choose between a P300 speller and a c-VEP BCI speller. At this stage, the system includes both online and offline applications for the P300 paradigm and only offline applications for the c-VEP paradigm.

The objective of this work was to start the development from scratch of a c-VEP system focusing the attention on the GUI design. This choice is motivated because, as discussed in Section 2.1, the GUI is the first aspect that needs to be investigated and built in order to achieve a high-performing BCI speller.

This study is divided in a sequence of steps:

— Design of the stimulus modulation for the interface, where multiple stimulation configurations can be chosen, in order to explore the influence of several stimulation parameters on the VEP response.

— Synchronization between the signal acquisition and the stimulus presentation to elicit the VEP response.
— Definition and realization of a pilot study to validate the designed system.

Stimulus modulation

Firstly, for the GUI design it is fundamental to consider many stimulation parameters as, for the moment, there is no universal set of optimal stimulus parameters, as discussed in Section 2.1.1.

In our work some stimulus parameters were set as default, and others were selectable from the visual stimulus module, shown in Figure 3.6. Indeed, the developed stimulus module allowed to customize the stimulus interface by specific stimulus type, choosing between the P300 and the c-VEP, and their respective parameters. However, in the following we will refer only to the c-VEP stimulus type.

The stimulus parameters set as default for the c-VEP stimulus type were:
— *Stimulation sequence*: binary m-sequence with a code length of 63 bits [Wei et al., 2016] and its time shifted versions by 2 bits.
— *Stimulation color*: if the bit in the corresponding binary stimulation sequence is 1 the character flickers in white, if it is 0 in black, as illustrated in Figure 3.10. The character in the virtual keyboard is placed in a circle with a dark grey background.

On the other hand, the stimulus parameters selectable from the stimulus module were:
— *Frequency of stimulation*: each element of the stimulation sequence can be flashed for 1, 2 or 4 frames. This means that on a 60 Hz screen, the characters on the virtual keyboard will thus flash at 60 Hz, 30 Hz or 15 Hz. Figure 3.8 shows an example of stimulus using different frequencies while keeping the same refresh rate of the screen and varying the number of frames.
— *Number of repetitions* of flashing sequence per target, also called *stimulus cycles*, that indicates how many times each target will flash for a complete stimulation sequence.
— *Keyboard arrangement*: two different virtual keyboard layouts, shown in Figure 3.7, can be enabled. The former respecting the principle of

equivalent neighbors, called “keyboard with borders”, the latter being a matrix of characters without additional flickering objects around the targets, called “keyboard without borders”.

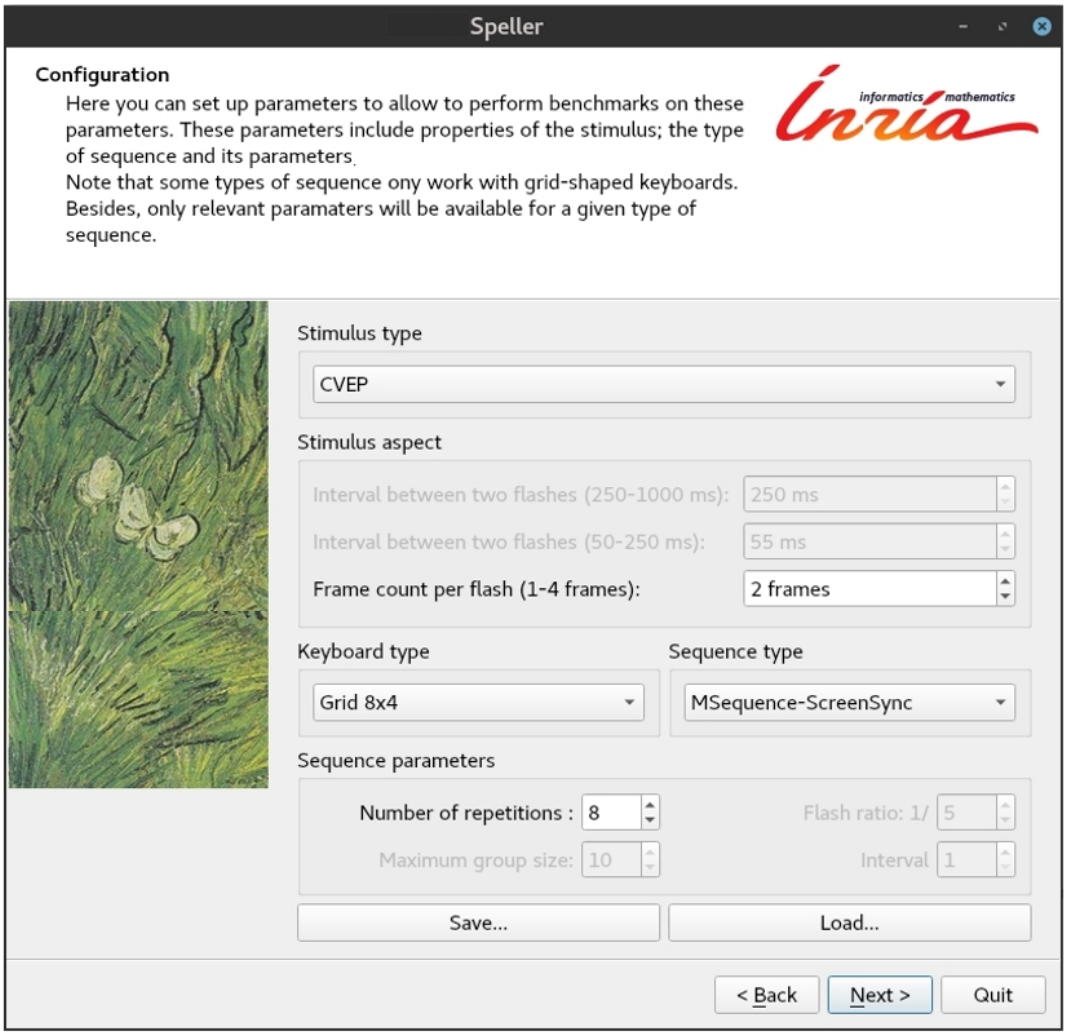


FIGURE 3.6 – Visual stimulus module for the c-VEP system. In black the stimulus parameters that can be set for the c-VEP stimulus type. In light grey the specific stimulation parameters of the P300 paradigm, that can be enabled when the stimulus type is “P300” (not considered in this section).

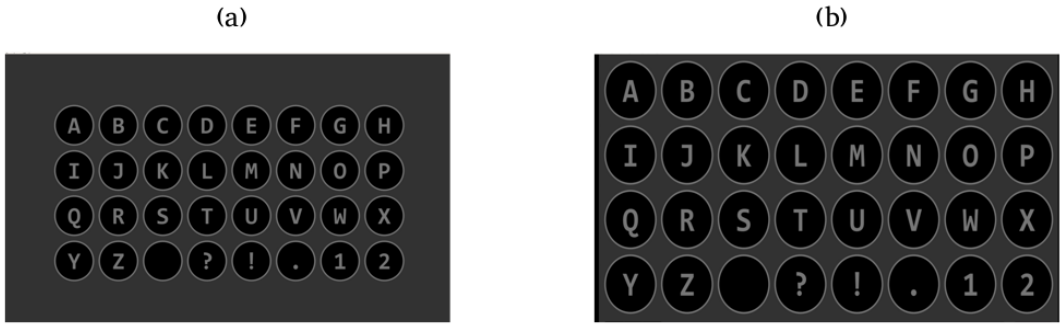


FIGURE 3.7 – The two layouts of the arrangement of characters for the designed virtual keyboard. (a) Keyboard with borders presents a layout with 32 targets (4×8 grid) surrounded by 28 non-targets, so that each target in the keyboard has eight neighbors. For each target, all neighbors keep fixed time lag relationships. (b) Keyboard without borders presents a layout with a 4×8 grid containing 32 characters: letters sorted alphabetically from A to Z followed by backspace, symbols " ? ", " ! ", " . " and numbers 1 and 2. The characters are the same in both keyboard arrangements.

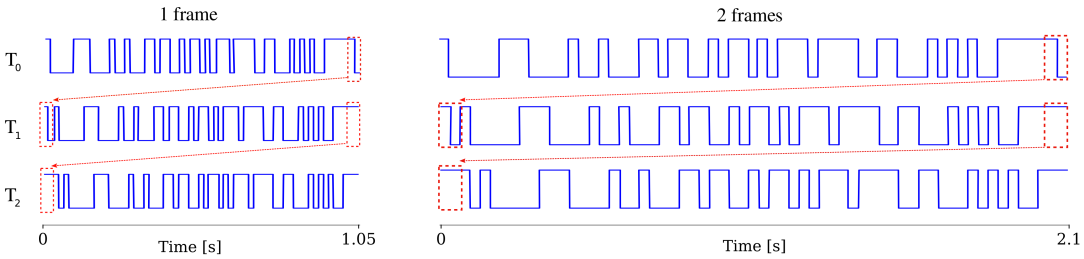


FIGURE 3.8 – Illustration of the circular-shift process for the stimulus sequence for the first 3 targets. On the left, the stimulus sequences have a frame rate of 60 Hz (1 bit = 1 frame) and on the right, a frame rate of 30 Hz (1 bit = 2 frames). The figure shows the stimulation sequence for the target T_0, for the target T_1, circularly shifted with a time lag τ_s with respect to T_0 and for the target T_2, circularly shifted with a time lag τ_s with respect to T_1. The red dash boxes indicate the time lag τ_s corresponding to the number of frames, listed in Table 3.3.

Synchronization

Another important aspect to consider in the implementation of a c-VEP speller is the synchronization between the EEG data recording and stimulus presentation. This aspect is fundamental to determinate accurate evoked responses in terms of timing.

In order to evaluate this synchronization we conceived a test using a photo-diode, which is a semiconductor device that converts light into an electrical current. The photo-diode was held in correspondence to the character "A" on the virtual keyboard, shown in Figure 3.7, which in turn was flashing following the 63-bit m-sequence (non shifted version), set as default parameter. A stimulus cycle was presented at the frequency stimulation of 60 Hz, since the refresh rate of the monitor was set at 60 Hz the length of each bit was of 1 frame, corresponding to 16.6 ms, and the total duration of the stimulus cycles was of 1.05 s.

The Figure 3.9 shows the coherency between the stimulation sequence and the signal recorded with the photo-diode: the amplitude is equal to zero when the binary sequence is 0 and reaches the maximum amplitude where the sequence is 1, as the character flickers in white.

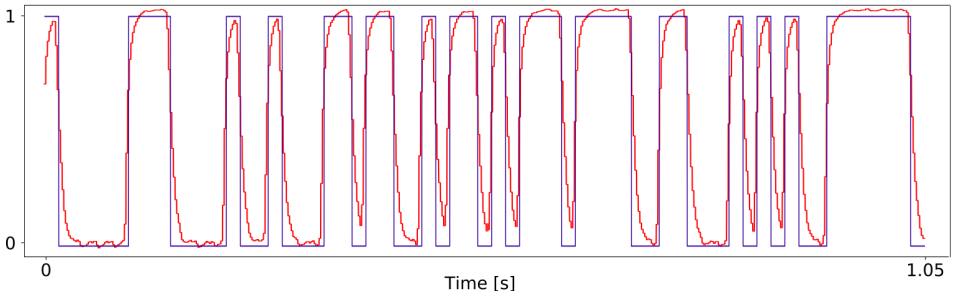


FIGURE 3.9 – Software-based synchronization between signal acquisition and stimulus presentation. In red the photo-diode signal recorded on the stimulus sequence (m-sequence) at 60 Hz (in blue). The plot shows the timing coherency between the stimulus sequence flashed by the system and the signal recorded by the photo-diode placed on the screen in correspondence of the character "A".

3.2.1 Pilot experiment

In order to validate the c-VEP system a pilot study was conducted. One subject participated voluntarily in the experiment with no previous experience with the c-VEP BCI. During the experiment the EEG signal of the subject was recorded from a ANT-Waveguard cap and a Refa8 amplifier (512 Hz sampling rate). Twelve electrodes were placed in positions Fz, F3, F4, Cz, C3, C4, Pz, P7, P8, Oz, O1, O2 and conductive gel was applied to these electrodes

and the ground (FPz) to set the impedance between the electrodes and the subject's skin below 10 kΩ.

The subject was seated in a comfortable armchair 100 cm away from the computer monitor with a refresh-rate of 60 Hz, placed in a quiet room. She was asked to spell the targets "A" and "R" with several experimental setups. Before starting the stimulation, each target was highlighted in blue on the virtual keyboard to indicate to the subject the target position. The characters were flashed following a 63-bit m-sequence and its time shifted version by 2 bits [Bin et al., 2011]. The stimulus parameters, in terms of number of frames and number stimulus cycles, followed the set of stimuli parameters reported in Table 3.3. Such parameters have been tested for both virtual keyboard layouts shown in Figure 3.7. Between each acquisition the subject was allowed to take a rest to avoid the effect of fatigue.

The stimuli are synchronized with the refresh rate of the monitor and a trigger signal is provided by a TCP network connection to synchronize the stimulus presentation and the EEG data recordings. Figure 3.10 gives an example of the presentation of the stimulation, illustrating the strategy for flashing at different frequencies of stimulation and the strategy of the trigger signal for the synchronization.

TABLE 3.3 – Stimulus parameters set applied for different experimental setups. The time lag $\tau_s = (2\ bits/60\ Hz) \cdot \#frames$. The duration of a single stimulation sequence t_s is computed as $t_s = (63\ bits/60\ Hz) \cdot \#frames$. The flashing duration of each target is $t_k = t_s \cdot \#stimulus\ cycles$

# frames	# stimulus cycles	τ_s	t_s	t_k
1	5	0.033 s	1.05 s	7.5 s
1	10	0.033 s	1.05 s	10.5 s
2	5	0.067 s	2.10 s	10.5 s
2	10	0.067 s	2.10 s	21.1 s
4	5	0.13 s	4.20 s	21 s
4	10	0.13 s	4.20 s	42 s

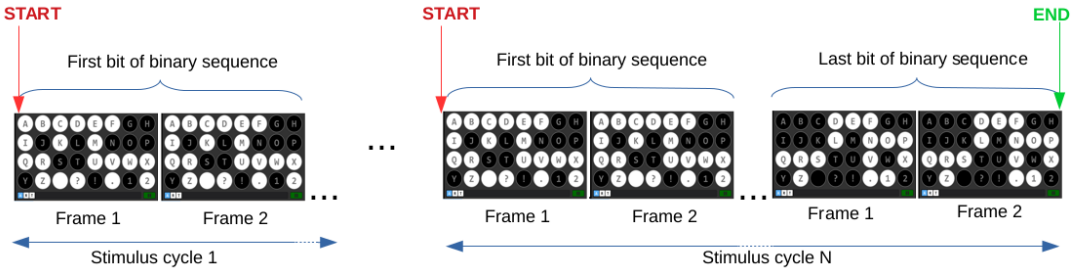


FIGURE 3.10 – Illustration of the stimulation strategy at a frequency of 30 Hz: each bit of the stimulation sequence of each character is flashed for two frames. In red the synchronization tags "START", sent at the beginning of each stimulus cycle, and in green the tag "END" sent at the end of the stimulus cycle N.

Offline processing

The EEG signal was filtered between 1 Hz and 20 Hz with a 5th order Butterworth bandpass filter. The recorded EEG signal was divided in epochs following the tags sent at the beginning of each flash sequence.

Then, in order to assess the repeatability of the VEP response for the same characters, the averaged correlation ρ^c_A of the VEP responses over the N stimulus cycles of the target "A", and the averaged correlation ρ^c_R of the VEP responses over the N stimulus cycles of the target "R" were computed.

Moreover the averaged cross-correlation ρ^c_{AR} between the N stimulus cycles of the VEP response of the target "A" and the VEP response of the target "R" was computed, in order to check the discriminability between two different targets in the recorded VEP responses. The expectation is a high value for ρ^c_A and ρ^c_R and a low value for ρ^c_{AR}.

The score λ^c evaluates the difference among ρ^c_{avg}, that is computed as the average between the auto-correlation ρ^c_A and ρ^c_R, and the cross-correlation ρ^c_{AR} per channel c for each parameters setting. The λ^c score is computed following equation (3.1), where std^c_{avg} is the average of the standard deviations of the VEP responses over the N stimulus cycles for the target "A" and for target "R", while std^c_{AR} is the standard deviation of the VEP responses over the N stimulus cycles of the target "A" and target "R".

$$\lambda^c = \frac{\rho^c_{avg} - \rho^c_{AR}}{std^c_{avg} + std^c_{AR}} \tag{3.1}$$

The score λ computed for each set of parameters is averaged over the three best channels, for which the average of ρ_A^c and ρ_R^c is the highest.

3.2.2 Results

Figure 3.12 shows the obtained curves of auto-correlation and cross-correlation and Figure 3.11 shows the λ scores computed for the different experimental setups. We can notice that using the keyboard layout without borders the values of λ obtained are higher than the ones obtained with the keyboard with borders, for all values of stimulation frequency and number of repetitions. The largest λ reached is 0.41 corresponding to 15 Hz and 5 repetitions. Also for 30 Hz the λ is higher for 5 repetitions than 10 repetitions. Instead for 60 Hz the λ is higher for 10 repetitions with a value equal to 0.37.

Analyzing the results reached for the keyboard with borders, the absolute minimum, equal to zero, is reached for 60 Hz and 10 repetitions. For 30 Hz the same value of λ is reached for both numbers of repetitions, while for 15 Hz the largest λ is achieved for 5 repetitions.

In Table 3.4 we can also find the average duration of the sequence of stimulation during the experiments t_{exp}. The results are computed as the interval between the tags sent at the beginning of each stimulation sequence for each stimulus cycle of a target. The results achieved are comparable with the results expected for t_s, listed in Table 3.3. This confirms the effectiveness of the synchronization method explained previously. However we can observe a small inaccuracy for the setups at 60 Hz, for both types of virtual keyboard.

3.2.3 Discussion and perspectives

The aim of this work was to investigate stimulus parameters that directly affect the stimulus duration, as shown in Table 3.3, to evaluate the impact on the repeatability of the VEP response and also the impact on the comfort of the subject, which are some fundamental aspects to consider when designing a GUI.

TABLE 3.4 – Results of the different experimental setups.

Frequency stimulation	Keyboard type	# repetitions	λ	t_{exp}	Best channels
60 Hz	with borders	5	0.088	1.047±0.015	C3
		10	0.004	1.048±0.015	-
	without borders	5	0.18±0.02	1.046±0.009	O1, Oz, O2
		10	0.24±0.02	1.048±0.013	O1, Oz, O2
30 Hz	with borders	5	0.13	2.09±0.02	P7
		10	0.13±0.03	2.094±0.015	O1, Oz, O2
	without borders	5	0.37±0.09	2.091±0.015	O1, Oz, O2
		10	0.30±0.02	2.1±0.02	O1, Oz, O2
15 Hz	with borders	5	0.21± 0.09	4.18±0.03	Cz, P3, P4
		10	0.16±0.06	4.19±0.02	O1, Oz, O2
	without borders	5	0.41 ±0.02	4.19±0.02	O1, Oz, O2
		10	0.31±0.01	4.19±0.02	O1, Oz, O2

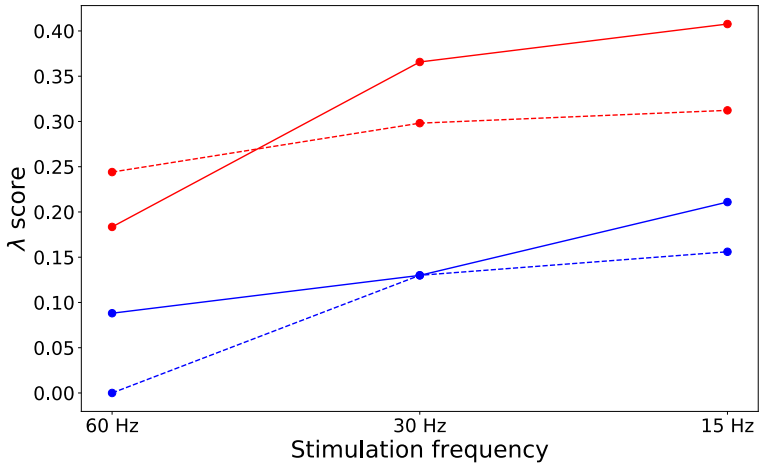


FIGURE 3.11 – The λ score for the different parameters of stimulation. In red λ score for the keyboard without borders and in blue the λ score for the keyboard with borders. Dash line indicates the number of repetitions equal to 10 and solid line the number of repetitions equal to 5.

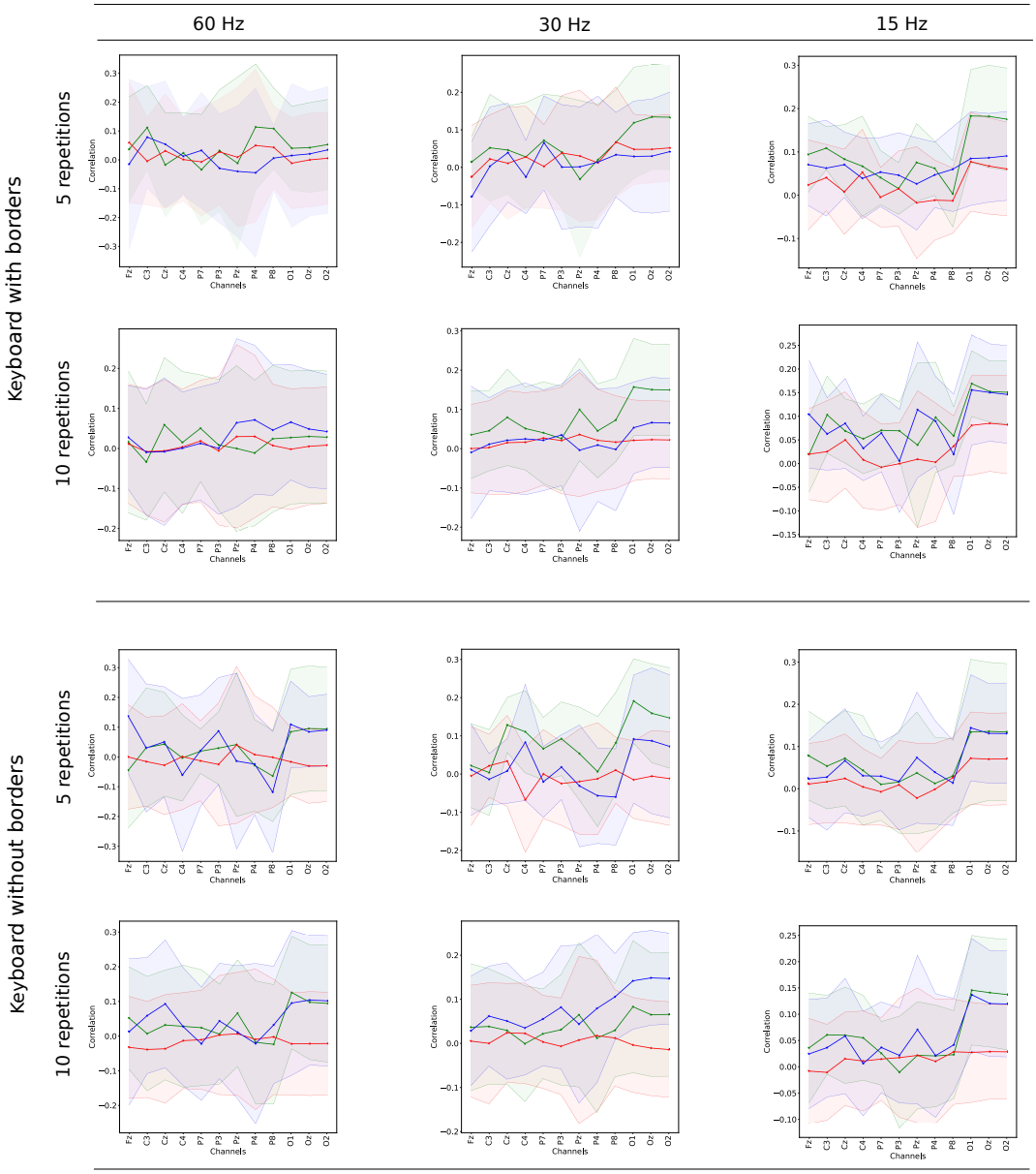


FIGURE 3.12 – Correlation curves of the tested stimulus parameters obtained for one subject during the spelling of the targets "A" and "R". The average auto-correlation per channel ρ_A in green and the average auto-correlation ρ_R per channel in blue. The average cross-correlation ρ_{AR} is represented in red.

Among the stimulus parameters we also selected the arrangement of the targets on the virtual keyboard, in order to evaluate the impact on the response and the possible additional differences between outer and inner targets. Therefore in the detailed experiment the chosen targets are either in a corner of the virtual keyboard (target "A") or in the middle (target "R"). In Figure 3.12 the auto-correlation curves are shown for both characters for different parameters. We can notice that, except for a few isolated cases such as for 30 Hz and the keyboard with borders, the auto-correlation curves (ρ^c_A and ρ^c_R) follow the same trend for both the targets, in particular on the best channels, listed in Table 3.4. So we can affirm that in our system the presence of flashes on the border does not impact measurably the VEP response, independently from the position of the target in the virtual keyboard.

Moreover the results in Table 3.4 show that the largest λ scores are reached for the keyboard without the borders, but there is not a predominant score corresponding to an unique value of stimulus frequency and number of repetitions. Analyzing the values in Figure 3.11 the obtained result is coherent with expectations. Indeed, with a higher frequency of stimulation the number of repetitions is higher, while better results are achieved with a lower number of repetitions when the value of stimulus frequency is lower. A possible explanation is that a slower stimulus, as in the case of 30 Hz and 15 Hz, requires a lower number of repetitions because if the stimulation becomes too long the subject could lose concentration.

Table 3.4 reports also the best channels for which the λ score is highest. It is interesting to notice that for many setups, and in particular for the setups with the highest λ values, the best channels are the occipital ones (O1, Oz, O2). This result could be expected because VEP responses are generated and recorded in the occipital part of the brain, as discussed in Section 2.1.

As our system has been designed from scratch, our results were compared with respect to a subset of a public dataset [Spüler et al., 2012], to prove its trustworthiness. The dataset will be analyzed in detail in Chapter 5. In this work the objective is to demonstrate that our results are comparable with the ones of the public dataset, capable of reaching high performance in a c-VEP speller. Therefore we computed λ scores for three subjects of the dataset, choosing the one that achieved the highest performance (subject AG001), one

that achieved good performance (AD001) and the one that achieved the lowest performance (subject AE001). The values of accuracy reported from literature [Spüler et al., 2012] are listed in Table 3.5. The average VEP response was computed over the same channels listed for our experiment, except for channels Oz replaced with channels PO8, and for the same number of repetitions, computing the offline spelling of the same target characters "A" and "R". However, the data was recorded for a keyboard with borders and 1 frame. In Table 3.5 we can find the results reached for the three subjects, they are comparable with respect to the results reached in our study, listed in Table 3.4 and shown in Figure 3.11, despite not having the same parameters of stimulation.

It is interesting to verify the presence of a comparable trend in the autocorrelation and cross-correlation curves under the standpoint of repeatability of VEP responses and discriminability between different characters, comparing for instance Figure 3.12 and Figure 3.13.

The most relevant difference between our results and the ones obtained for the dataset of [Spüler et al., 2012] is that high values of λ were achieved with a keyboard with borders for the dataset [Spüler et al., 2012]. Instead in our case, as evident in Figure 3.11, the score reached with the keyboard with borders is lower that the score reached with the other keyboard without borders. One main reason of this result is that our subject confirmed the difficult and discomfort during the acquisition with the keyboard with borders, for two main reasons. Firstly, the targets in the keyboard are smaller due to the presence of the borders and, secondly, the flashing on the borders caused for the subject an element of distraction and annoyance.

This shows once again that there is no universal set of parameters but most of the performance of a system depends on the implementation of the GUI, on the experimental setup and on the subject itself.

However in Table 3.5 are listed the values of accuracy reported in the literature for the dataset analyzed [Spüler et al., 2012]. It is interesting to notice that largest values of λ score correspond to higher values of accuracy. This opens the way for the use of λ as a possible parameter to evaluate the performance of a subject, as will be discussed in more detail in the following chapter.

Finally, this preliminary study allowed to develop a GUI effective for a c-VEP speller, as confirmed by the results obtained, comparable with other data available in literature. It allowed as well to define a set of parameters to be used for future experiments. Moreover it confirmed the variability of the responses while changing the stimulation parameters and how much it is necessary to develop systems that can cope well with such variability between subjects. A possible solution to this issue will be presented in the following chapters.

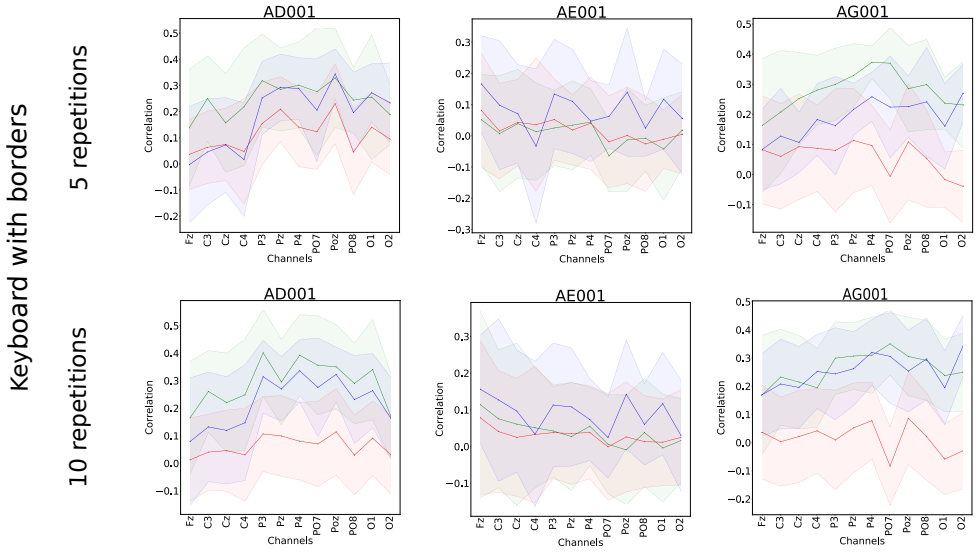


FIGURE 3.13 – Correlation curves obtained for the offline spelling of the target "A" and "R" for three subjects belonging to the dataset [Spüler et al., 2012]. The average auto-correlation per channel ρ_A in green and the average auto-correlation ρ_R per channel in blue. The average cross-correlation ρ_{AR} is represented in red.

TABLE 3.5 – λ score reached for three subjects of the public dataset with accuracy values reported from literature [Spüler et al., 2012].

Subject	λ over 5 repetitions	λ over 10 repetitions	accuracy
AD001	0.34 ± 0.2	0.68 ± 0.11	99.13 %
AE001	0.09 ± 0.1	0.13 ± 0.06	77 %
AG001	0.63 ± 0.2	0.9 ± 0.21	100 %

3.3 Conclusions

In this chapter two studies were presented based on two different BCI speller systems, both addressing the analysis of some stimulation parameters.

The first study was based on the P300 paradigm with the aim of improving the discriminability using a letter-based flashing strategy. After the implementation, some preliminary experiments have shown that the improvement in discriminability comes at the expense of the flashing duration, whose extension can compromise the overall performance of the speller and can also negatively impact the comfort of the subject.

The second study was based on the design of a c-VEP BCI speller, in which it is still possible to maintain a high discriminability and at the same time a good performance. We focused the research on the GUI as well as on some stimulus parameters, and we realized a pilot study, also compared with data present in literature. Finally we can affirm that there is no universal set of stimulation parameters available for a c-VEP system, but the system should be able to adapt to each subject.

4 Adaptive BCI Systems

In Chapter 2 we reviewed two BCI applications: the BCI speller and the MI-BCI system and we concluded that, in order to tackle some limitations of these BCI systems, it is fundamental to develop an user-centered system to increase the BCI performance. In this chapter two different works concerning the realization of an user-centered system will be presented.

The former deals with the development of an adaptive c-VEP BCI speller, improving the work presented in Chapter 3. There we proposed two studies to evaluate the influence of some stimulus parameters for a P300 speller and for a c-VEP speller. About the c-VEP BCI, we focused our research on the GUI as well as on some stimulus parameters. The pilot study showed the variability of the responses caused by different stimulation parameters and how much it is necessary to develop systems that can cope well with such variability between subject.

In the first part of this chapter we propose a possible solution to this issue by an adaptive parameter setting c-VEP BCI system, in which the traditional calibration phase is replaced by a shorter adaptive setting phase, in order to define the most pleasant and best performing stimulus for each subject across sessions.

The second work presented in this chapter concerns the realization of a MI-BCI system for a disabled user in the context of a live BCI competition. In particular we present the long training necessary to obtain a performing system and the competition, highlighting that in this type of system the focus must be on the user.

4.1 Adaptive parameter setting in c-VEP BCI

As explained in Chapter 2, many studies investigate the effect of stimulus parameters on the target flashing modulation. Analyzing these studies shows clearly that it is not possible to define an universal optimal stimulus parameter setting suitable for each BCI user. In order to obtain a BCI system with high communication speed, that respects the subject's comfort, it is necessary to develop a system adaptable to the subject.

Moreover, the pilot experiment presented in Chapter 3 confirms this observation and paves the way for the realization of an adaptive c-VEP system that will be presented in this section. Thanks to this study we defined a set of stimulus parameters to deploy in our system and we chose to investigate the impact and the variability of the stimulus presentation rate. Hence, in our study we developed a subject-dependent system with four different stimulus presentation rates of 15 Hz, 20 Hz, 30 Hz and 60 Hz. The objective is to find the optimal presentation frequency to obtain the most suitable stimulus per each subject. We demonstrate that the decreasing of the stimulation frequency does not imply a decrease of the system performance, showing the importance of the BCI-user adaptation.

4.1.1 Adaptive parameter setting phase

In a traditional c-VEP BCI with a refresh rate of 60 Hz and target encoded by binary sequence, each element of the sequence is flashed on the screen for a time $t_b = 16.67\ ms$, corresponding to the duration of one frame. We developed a system in which the targets are encoded by a 63-bit m-sequence, but each element of this sequence can be flashed for 1, 2, 3 or 4 frames. On a 60-Hz screen, the characters on the virtual keyboard will flash at 60 Hz, 30 Hz, 20 Hz or 15 Hz, as explained in Section 3.2. In this way it is possible to flash the target faster or slower and to find the stimulus pattern most adapted to each subject, while maintaining a system with high performance.

Our protocol thus starts with an adaptive parameter setting phase. In this phase the subject has to focus on the targets of a word of eight characters, shown in Figure 4.1, so the word can be divided in four pairs of targets and each pair follows the setting reported in Table 4.1. The flashing rate is thus

changed every two targets. We compare the correlation between all the VEP responses recorded for each stimulus cycle of the two targets, flashed with the same frame value per each channel c, to detect for which stimulus setting the evoked response of the subject is most prominent. Let ρ_1^c be the averaged correlation of the VEP responses over the N stimulus cycles of the target 1, ρ_2^c the averaged correlation of the VEP responses over the N stimulus cycles of the target 2 and ρ_{12}^c the averaged cross-correlation between the N stimulus cycles of the VEP response of the target 1 and the VEP response of the target 2. Based on the fact that when the user is gazing at a target, the specific VEP recorded in the EEG should be the same for each stimulus cycle of the same target, the expectation is a high value for ρ_1^c and ρ_2^c and a low value for ρ_{12}^c.

The score λ^c, introduced in Section 3.2, evaluates the difference between ρ_{avg}^c, that is computed as the average between the auto-correlation ρ_1^c and ρ_2^c, and the cross-correlation ρ_{12}^c per channel c for each parameter setting. The λ^c score is computed following formula (4.1), where std_{avg}^c is the average of the standard deviations of the VEP responses over the N stimulus cycles for target 1 and target 2, while std_{12}^c is the standard deviation of the cross-correlation ρ_{12}^c.

$$\lambda^c = \frac{\rho_{avg}^c - \rho_{12}^c}{std_{avg}^c + std_{12}^c} \tag{4.1}$$

At the end of the data acquisition of this adaptive phase the score λ^c is averaged over the best three channels, for which ρ_{avg}^c is the highest. The largest score λ is chosen to select the best stimulation parameter.

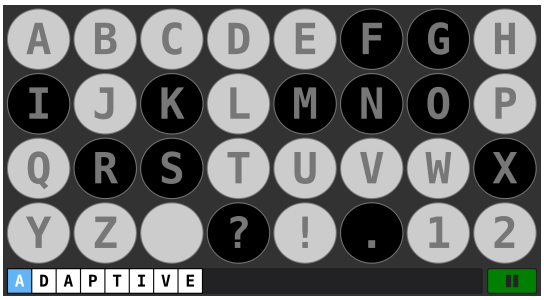


FIGURE 4.1 – Virtual keyboard developed for the c-VEP BCI system. During the stimulation the target is highlighted in blue below the keyboard.

TABLE 4.1 – Stimulus parameters set for each two consecutive targets in the adaptive parameter setting phase. The time lag is $\tau_s = (2\ bits/60\ Hz) \cdot \#frames$. The length of one sequence t_s is computed as $t_s = (63\ bits/60\ Hz) \cdot \#frames$. The flashing duration of each target is $t_k = t_s \cdot \#stimulus\ cycles$

target1-target2	# frames	# stim. cycl.	τ_s	t_s	t_k
A - D	1	10	0.033 s	1.05 s	10.50 s
A - P	2	8	0.067 s	2.10 s	16.80 s
T - I	3	5	0.10 s	3.15 s	15.75 s
V - E	4	4	0.13 s	4.20 s	16.80 s

4.1.2 Experimental setup

The BCI software consists of OpenViBE [Renard et al., 2010] for signal acquisition, and a custom keyboard-display control software, which we developed in C++. This software runs on a Windows 7 computer with an Intel(R) Xeon(R) processor. Two different LCD monitors are used: one (DELL U2711) is used during the acquisition to monitor the EEG signal quality, and the other (DELL 2709W) is set at 60 Hz with a resolution of 1920×1080 with a NVIDIA Quadro FX 580 graphic card and it is used for the stimuli presentation on a virtual keyboard. The keyboard, displayed in Figure 4.1, is a 4×8 matrix containing 32 characters: letters sorted alphabetically from A to Z followed by space, symbols "?", "!", "." and numbers 1 and 2. Each character is placed in a light grey circle over a dark grey background. Below the matrix of targets a text field shows the characters of the word that the subject has to gaze at. Each character of the virtual keyboard flashes according to a binary sequence composed of 0 and 1. If the bit in the corresponding binary stimulation sequence is 1 the character flickers in light grey, if it is 0 in black, as illustrated in Figure 4.1. The color combination light grey/dark grey was chosen instead of white/black, in order to make the contrast more comfortable for the subjects [Gembler et al., 2020; Bieger et al., 2010]. The stimuli are synchronized with the refresh rate of the monitor and a trigger signal is provided by a TCP network connection to synchronize the stimulus presentation and the EEG data recordings.

FIGURE 4.2 – Experimental setup. A subject wearing an EEG headset during the second phase of the experimental protocol. The subject is focusing his attention on the character highlighted in blue in the word written below the virtual keyboard. The other screen on the left is used to monitor in real time the subject's EEG.

Participants and data acquisition

Nine healthy volunteers participated in the c-VEP BCI experiment. The experiment took place in our premises at Inria and was approved by the Operational Committee for the assessment of Legal and Ethical risks of the institute. All subjects had normal or corrected to normal vision and did not suffer from epilepsy or any other nervous disease. Each subject took part in two identical sessions, half a week apart, and all of them completed the whole experiment. All the subjects were c-VEP BCI-naive participants, except one as listed in Table 4.2.

During the experiment, each subject was seated in a comfortable armchair 100 cm away from the computer monitor placed in a quiet room. During each session the EEG signal of the subject was recorded from ANT-Waveguard cap and a Refa8 amplifier (512 Hz sampling rate). To set the impedance between the electrodes and the subject's skin below 10 kΩ, a conductive gel was applied to the ground (FPz) and to the 12 electrodes placed in positions Fz, F3, F4, Cz, C3, C4, Pz, P7, P8, Oz, O1, O2. A picture of the experimental setup is shown in Figure 4.2.

TABLE 4.2 – Information about subjects that participated to the experiments.

Subject	Sex	Previous BCI experience	Problem during c-VEP experience
S1	M	P300	Difficulty of concentration in Session 2, problem in Fz
S2	F	P300	
S3	F	P300; c-VEP	
S4	F	-	
S5	M	-	
S6	M	P300	Difficulty to focus on central characters in Session 1
S7	M	-	Problem in O2
S8	F	P300	
S9	M	-	Difficulty of concentration in Session 2

Each session consisted in 2 phases: the adaptive setting phase and a second phase in which the subject focuses on imposed characters, in which each target flashes according to the set of parameters found during the adaptive phase. To avoid the effect of fatigue on the experiment, the subject was allowed to take a 5-minute rest between the two phases. Each session lasted around 45–60 minutes, including the time for the experiment preparation (positioning of the cap and conductive gel application) and the time for the data acquisition.

During the adaptive setting phase a word of eight characters is displayed on the screen, below the keyboard. The characters are flashed following a 63-bit m-sequence and its time shifted version by 2 bits [Bin et al., 2011], an example for the circular shift of the stimulus sequence can be seen in Figure 3.8, in the previous chapter. The stimulus parameters, in terms of number of frames and number of stimulus cycles, follow the settings reported in Table 4.1. Each target, before starting the stimulation, is highlighted in blue on the virtual keyboard to indicate to the subject the target position. This adaptive setting phase lasts around 2 minutes.

At the end of this phase the data collected is processed, as explained in Subsection 4.1.1, and the output of the processing gives the best set of parameters per subject for that session. Moreover, the subject was asked by an oral question at the end of the adaptive parameter setting phase for which value of frame rate the visual stimulus was more comfortable. Then, during the second phase, the subject had to focus his/her attention on the characters of a specific word written below the keyboard. In this phase each target was highlighted in blue before starting flashing, then all the characters flashed following the set of best parameters computed during the first phase. Each subject had to spell five different words (42 characters in total and 10 targets), with a pause of one minute between each word.

Offline processing

The EEG data $\mathbf{X}$ was collected with N stimulus cycles on c channels, bandpass filtered between 4 and 22 Hz with a Butterworth filter of order 4 and notch filter at 50 Hz. The canonical correlation analysis (CCA) [Spüler et al., 2014] was applied as spatial filter to improve the signal-to-noise ratio of the EEG signal (see Section 2.1). The electrodes that proved an high interference or problem during the acquisition phase were not considered in subsequent processing. To compute $\mathbf{S}$ we average over the number of stimulus cycles N the responses recorded for the first character of the first word that the subject had to gaze during the second phase. We considered only the three best channels selected during the adaptive setting phase and then replicated N times the signal to obtain $\mathbf{S}$ [Spüler et al., 2012], in this way $\mathbf{X}$ and $\mathbf{S}$ have the same number of stimulus cycles N. Then W_X is multiplied with $\mathbf{X}$ to compute the spatially filtered signal x. For target identification the method of template matching [Bin et al., 2011] is used. The reference template T_0 is calculated by averaging the signal x of the first character of the first word over the N stimulus cycles. The templates of all other targets T_k ($k = 0, ..., 31$) are generated by circularly shifting the template T_0. The duration t_s of the template and the time lag τ_s between two consecutive targets depend on the number of frames set for each subject, as listed in Table 4.1. To detect the attended target we compute a cumulative correlation per target. We segment the spatially filtered signal in epochs starting at the "start" trigger, sent at the

beginning of each stimulus cycle, and lasting the length of a stimulation sequence, specific to each subject (see Table 4.4), in order to obtain the epochs x_n, with $n = 0, ..., N$. To set the N stimulus cycles, the cumulative correlation between the stimulus repetitions is computed and an arbitrary threshold is fixed at 0.8, considering the normalized correlation of each subject. Finally we compute the cumulative correlation coefficient ρ_k between each template T_k and the epochs x_n, following the formula (4.2).

$$\rho_k = \sum_{n=0}^{N} \frac{T_k \cdot x_n}{\sqrt{T_k \cdot T_k} \cdot \sqrt{x_n \cdot x_n}} \tag{4.2}$$

The target k with the largest coefficient ρ_k is detected as the attended target k_{target}. The offline spelling accuracy of each word is computed to evaluate the performance of the BCI system. The number of correctly detected characters is computed for each word that the subject had to spell during the second phase. Finally the accuracy per word is computed as the proportion of correctly detected characters.

4.1.3 Results

Table 4.3 lists the λ scores obtained per subject and session as well as the frequency rate preferred by the subject. We can notice that the highest obtained value of λ corresponds to the frequency rate that each subject expressed as the most comfortable, except for subjects S1 and S6 that did not express a preference. Moreover, we can observe that only for the subjects S1, S8 and S9 the optimal frame rate is unchanged between the two sessions. For the other subjects the optimal number of frames changes between one session and the other. Across all subjects, subject S2 reached the largest λ at the end of the first session equal to 1.33 ± 0.32, corresponding to 2 frames. If we compare this result with respect to the results obtained at the end of the second session, we can notice that for the same number of frames the λ value is 0.64 ± 0.12, but the largest λ is reached for 4 frames, with a value equal to 1.46 ± 0.12. The same trend can be observed for subject S6 from 0.56 with 4 frames and 0.71 with 3 frames. It is interesting to notice that for some subjects the difference of λ values is consistent across the different number of frames. For instance, subject S5 reached a λ equal to 1.21 with 4 frames

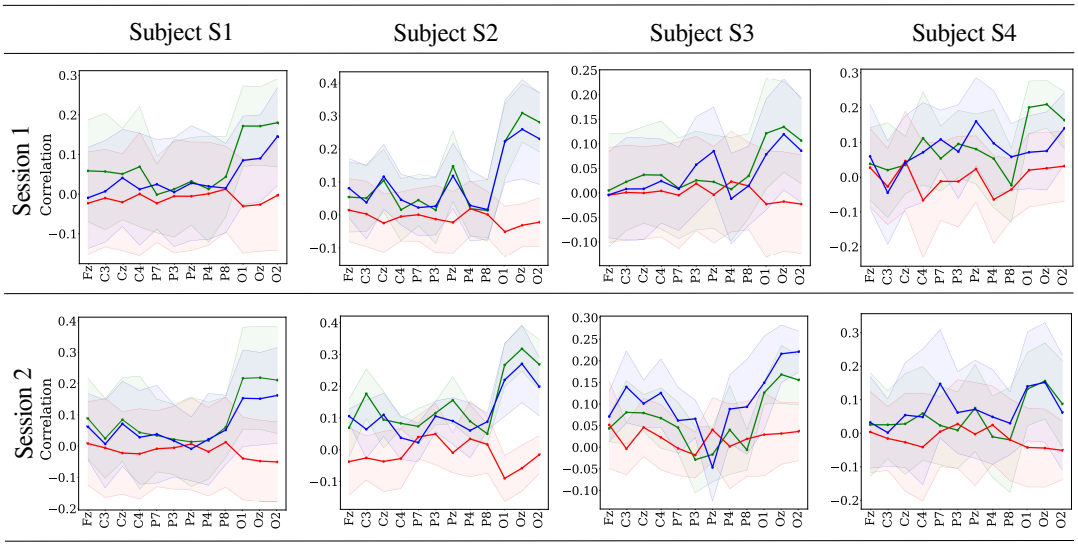


FIGURE 4.3 – Correlation curves of the optimal number of frames, listed in Table 4.4, obtained for a subset of subjects at the end of Session 1 and Session 2. The average auto-correlation per channel ρ_1 in green and the average auto-correlation ρ_2 per channel in blue. The average cross-correlation $\rho 12$ is represented in red.

and equal to 0.37 with 1 frame in Session 1. The same for subject S9 with a λ around 0.70 with 3 frames and around 0.04 with 1 frame, for both sessions.

Figure 4.3 shows the auto-correlation and cross-correlation curves obtained for a subset of subjects for the best frame rate selected at the end of the adaptive parameter setting phase. The parameters obtained at the end of the adaptive setting phase and set for the second phase are listed in Table 4.4. In Figure 4.4 we show the averaged offline spelling accuracy, over the five words, of each subject for each session. The accuracy of subject S2 reaches a good level, 96 % in the second session and over 82 % in the first one. We can notice that for subjects S2, S3, S5, S6 and S8 there is an increase of performance during the second session. On the contrary for subjects S1, S4, S7 and S9 the performance decreased from Session 1 to Session 2. For instance, the performance during the first session of subject S1, reaching 67 % of accuracy, is better than the one of his second session, with an accuracy of around 15 %.

TABLE 4.3 – Average λ score with standard deviation and the rate preferred frame rate during the adaptive parameter setting phase in Session 1 and Session 2. In bold the maximum λ score corresponding to the frame rate selected at the end of the adaptive parameter setting phase.

Session 1					
Subject	1 frame	2 frames	3 frames	4 frames	Subject's preference
S1	0.24 ± 0.02	$\mathbf{0.40 \pm 0.02}$	0.29 ± 0.05	0.11 ± 0.3	no preference
S2	1.23 ± 0.12	$\mathbf{1.33 \pm 0.32}$	1.10 ± 0.18	1.03 ± 0.08	2 frames
S3	0.43 ± 0.14	$\mathbf{0.71 \pm 0.06}$	0.51 ± 0.03	0.34 ± 0.05	2 frames
S4	0.29 ± 0.02	0.33 ± 0.03	$\mathbf{0.58 \pm 0.03}$	0.57 ± 0.13	3 frames
S5	0.37 ± 0.06	$0.54 \pm 0.0.8$	0.60 ± 0.03	$\mathbf{1.21 \pm 0.09}$	no preference
S6	0.21 ± 0.04	0.24 ± 0.09	0.28 ± 0.03	$\mathbf{0.56 \pm 0.08}$	2 frames
S7	0.15 ± 0.01	0.16 ± 0.06	0.48 ± 0.03	$\mathbf{0.67 \pm 0.04}$	$3-4$ frames
S8	0.30 ± 0.03	0.27 ± 0.02	0.38 ± 0.05	$\mathbf{0.82 \pm 0.01}$	$3-4$ frames
S9	0.04 ± 0.02	0.31 ± 0.03	$\mathbf{0.70 \pm 0.01}$	0.44 ± 0.01	3 frames

Session 2					
Subject	1 frame	2 frames	3 frames	4 frames	Subject's preference
S1	0.24 ± 0.02	$\mathbf{0.45 \pm 0.02}$	0.42 ± 0.05	0.01 ± 0.3	no preference
S2	0.36 ± 0.01	0.64 ± 0.02	0.87 ± 0.06	$\mathbf{1.46 \pm 0.17}$	4 frames
S3	0.14 ± 0.03	0.61 ± 0.01	0.34 ± 0.08	$\mathbf{0.63 \pm 0.08}$	4 frames
S4	0.3 ± 0.02	$\mathbf{0.44 \pm 0.04}$	0.42 ± 0.05	0.13 ± 0.08	2 frames
S5	0.29 ± 0.04	0.44 ± 0.07	$\mathbf{0.76 \pm 0.07}$	0.44 ± 0.5	4 frames
S6	0.39 ± 0.05	0.53 ± 0.01	$\mathbf{0.71 \pm 0.1}$	0.48 ± 0.06	no preference
S7	0.01 ± 0.03	0.41 ± 0.01	$\mathbf{0.54 \pm 0.01}$	0.52 ± 0.12	$3-4$ frames
S8	0.54 ± 0.02	0.55 ± 0.02	0.71 ± 0.03	$\mathbf{0.99 \pm 0.07}$	$3-4$ frames
S9	0.04 ± 0.05	0.16 ± 0.06	$\mathbf{0.74 \pm 0.01}$	0.28 ± 0.02	3 frames

TABLE 4.4 – Summary of the adaptive stimuli parameter setting. Number of frames, number of stimulus cycles (st.cyc.) and stimulus duration of the target (t_k) per subject and across sessions.

Session 1				**Session 2**			
subject	# frames	# st.cyc.	$t_k[s]$	subject	# frames	# st.cyc.	$t_k[s]$
S1	2	7	14.7	S1	2	7	14.7
S2	2	7	14.7	S2	4	3	12.6
S3	2	6	12.6	S3	4	3	12.6
S4	3	4	12.6	S4	2	6	12.6
S5	4	3	12.6	S5	3	4	12.6
S6	4	3	12.6	S6	3	4	12.6
S7	4	3	12.6	S7	3	4	12.6
S8	4	3	12.6	S8	4	3	12.6
S9	3	4	12.6	S9	3	4	12.6

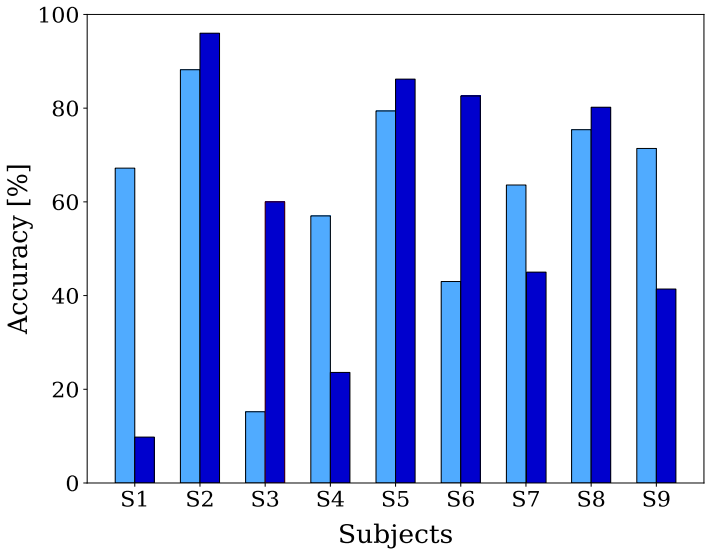


FIGURE 4.4 – The box plots report the averaged accuracy over five words obtained in the offline spelling for each subject and each session. The Session 1 is represented in light blue, Session 2 in dark blue.

4.1.4 Discussion

In our study we explored the influence of frequency stimulation on the VEP responses by the development of an adaptive setting phase. At the end of this adaptive phase we computed the λ score that could be considered as a performance predictor for a c-VEP BCI system. Indeed if we compare the largest λ value with respect to the accuracy values reached in the offline spelling, shown in Figure 4.5, we can notice that a large λ value corresponds to higher accuracy value for 6 subjects out of 9, with the exception of subjects S1, S4 and S5. The obtained results show an important variability of the stimulation frequency, both among subjects and across sessions. Indeed one of the challenges of using BCIs over extended periods of time is the variation of the user's performance from a session to another. There are many factors that can influence the BCI user's performance between different sessions, for example distraction, visual fatigue, loss of concentration, motivation and comfort [Kleih and Kübler, 2015]. The objective of the proposed method is to obtain a system in which it is possible to define the more comfortable stimulus for each subject, in order to increase the performance of the system, replacing the traditional calibration with a shorter adaptive setting phase. Table 4.3 shows that the most comfortable frame rate expressed by the subject is in most cases the one for which the performance is the highest. This can explain why a subject that performs well for a frequency rate performed less well for another frequency rate during the same session, demonstrating the need to set the parameters of stimulation for each session.

Moreover it is interesting to notice that none of the subjects conducted the second phase of the experiment with a frequency of 60 Hz. As illustrated in Table 4.4, the flashing duration of each target is not longer even when the frame rate is longer than one frame, which is the most common value of frame rate (corresponding to 60 Hz) used in c-VEP systems, as explained in Section 2.1.1. Indeed the decrease of the flashing pattern frequency implies an increase of the duration of a stimulus cycle, but we compensated the total duration of the stimulation by setting a lower number of stimulus cycles. This means that the stimulus duration at different frame rates does not impact the performance of our system.

Comparing the values of frequency stimulation set in our system with the

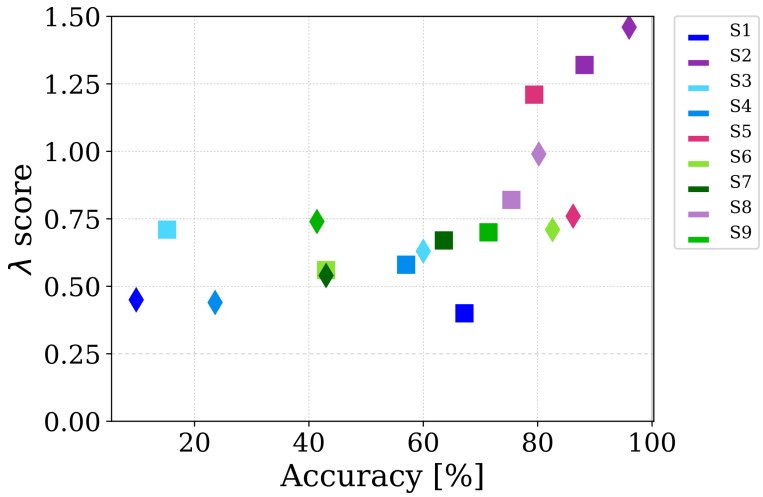


FIGURE 4.5 – The best λ scores, in bold in Table 4.3, of each subject with respect to the offline spelling accuracy, showed in Figure 4.4. Session 1 is represented with a square, Session 2 with a diamond. The λ score increases according to the accuracy.

work proposed by [Borhanazad et al., 2019], they reported that many subjects found more comfortable a visual stimulation with higher frequencies (60 Hz, 90 Hz and 120 Hz) than with a lower frequency equal to 40 Hz, that for some subjects was classified as "very uncomfortable". On the contrary, in our study many subjects reached a good performance with a frequency of stimulation of 15 Hz or 20 Hz, that corresponds also to their visual stimulation preference. This confirms the difficulty to define an universal set of stimulus parameters, because each system has its stimulus design.

It would certainly be interesting to understand why some subjects reach a high accuracy value and others do not, achieving an average accuracy lower with respect to other c-VEP BCI systems [Bin et al., 2011; Spüler et al., 2012]. Among many reasons that can explain this difference, there is the quality of the signal recorded during the experiments. Figure 4.6 shows the spatially filtered responses for participants with the worst and the best performance in Session 2, in terms of accuracy (subject S1 and subject S2 respectively). Comparing the VEP responses in Figure 4.6 is evident that there is a better repeatability of the VEP response for subject S2, for whom the accuracy is around 80%, and a larger variability on the responses for subject S1, who obtains an accuracy of around 15%. Lack of repeatability can be due to the inexperience of the subject, or the discomfort of the subject during the session, or to environmental noise.

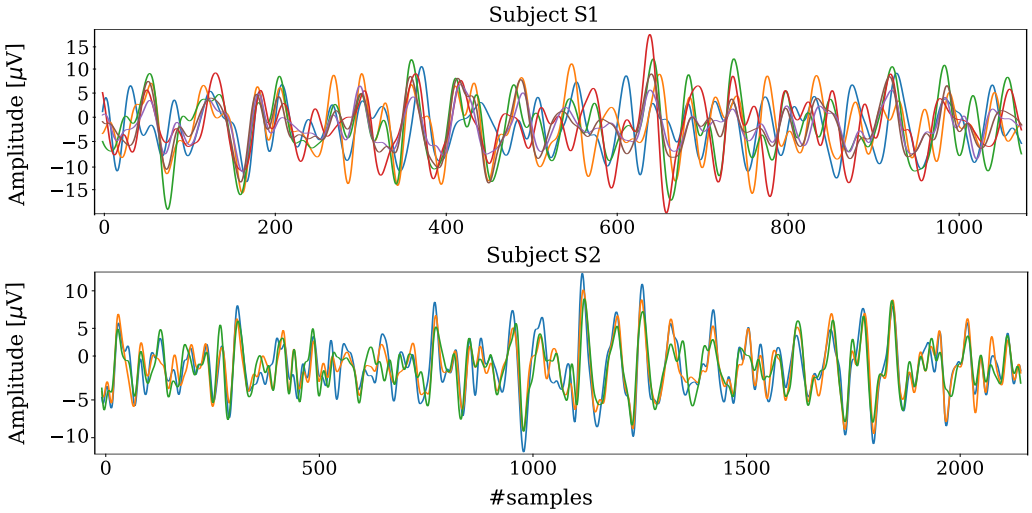


FIGURE 4.6 – The spatially filtered VEP responses recorded at each stimulus cycle overlapped for target k, acquired during Session 2 for subject S1 and subject S2, respectively the subjects that perform worst and best.

In fact, S1 reported to be really tired with difficulty of concentration during Session 2, as reported in Table 4.2. We can observe the same trend also for subject S9, who reported the same feeling in Session 2, and as consequence we can observe in Figure 4.4 a decrease of performance in Session 2 with respect to Session 1.

Another factor that can impact the performance can be the difficulty to focus the attention on the virtual keyboard, especially for subjects that do not have previous experience with the system. For instance, S6 declared it was difficult to focus his concentration during Session 1 on central characters in the arrangement of the virtual keyboard, due to the adjacent characters. However, during Session 2 he acquired more confidence with the system and did not report any discomfort due to the visual stimulation.

In the future design of an online c-VEP BCI with an adaptive parameter setting phase, this phase can effectively replace the traditional calibration phase. Indeed the VEP responses evoked by the optimal set of stimulus parameters during the adaptive parameter setting phase can be applied to train the classifier. Moreover many other improvements can be done for such an online c-VEP BCI, like the inclusion of other variables in the adaptive parameter setting phase, such as the selection of the stimulus sequence, stimulation color, and development of an online early stopping method to find the optimal

number of stimulus cycles per subject. Moreover, to further improve our system, the next step will be to identify the disturbance factors and find methods to remove them in the VEP responses, by different applications of spatial filters, or by modeling the VEP response and the external disturbances and noise.

To conclude, in this work we parameterized the stimulus modulation with four different stimulus presentation rates. We developed an experimental protocol that deploys a preliminary phase to define the optimal setting of stimulation frequency, tackling the problem of variability inter-subject and intra-subject. Optimally, these results could help with the design of a subject-dependent c-VEP BCI with high communication performance.

4.2 Design of MI-BCI system for a disabled user

In this section we present the design of MI-BCI combining mental imaginary and cognitive tasks for a severely motor impaired user, involved in the BCI race of the Cybathlon event [Riener, 2016]. The development of a high-performance MI-BCI requires a long training, in the order of weeks and months, as explained in Section 2. We present all the procedures followed to realize an effective MI-BCI, from the first contact of the user with a BCI technology to actually playing a video-game by her EEG. We defined a multi-stage user-centered training protocol in order to successfully control a BCI, even in a stressful situation, such as that of a competition.

The Cybathlon [Riener, 2016] is a competition of people with severe motor disability who, thanks to the use of assistive technology, can compete in different disciplines, such as the BCI-race. In the BCI-race the user of the system becomes the pilot in a virtual race game against up to three other pilots, in which each pilot has to control his/her virtual car by his/her mental tasks. The virtual car moves on the race track, and if the control command is correct it receives a speed boost, otherwise it is penalized, proceeding more slowly. The car can be controlled by four different commands (go straight, go right, go left and turn on the lights) and the pilot can assign his/her specific mental task to each command, without any restriction.

To efficiently train our pilot we deployed a multi-stage training, that consisted in an investigation phase to detect the subject-dependent specific mental and cognitive tasks, followed by a training phase.

4.2.1 Investigation phase

The investigation phase is fundamental to define the MI tasks most suitable for the subject. Indeed to select the mental tasks multiple criteria must be fulfilled: the subject is able to perform them, the individual mental task has to produce a recognizable brain pattern and they must not cause unwanted side effects, like spams, discomfort or stress.

We collected data over several sessions in one month, from the middle of June 2019 to the middle of July 2019. This phase took time because this experience was new both for the user, who had never used a BCI system before, and for our team, indeed for the first time we worked with a disabled person, which obviously requires specific attention. Therefore, a preliminary phase was necessary to create collaborative relationship between the team and the user, to allow the user to become more familiar with the hardware and also to allow the team to understand how to effectively manage this type of experience, defining a suitable experimental protocol.

Our pilot is a 32 years old woman, with a neurodegenerative disease since she was 7 years old. She has no cognitive disability but severe motor disabilities. Moreover, she participated to different sport competitions for disabled people but she did not have any experience with BCI, as said before.

The experiments took place in a room located in the pilot's living center "Centre René Labreuille" in Le Cannet. During each session the EEG signal of the subject was recorded from a ANT-Waveguard cap with a Refa8 amplifier (512 Hz sampling rate). To set the impedance between the electrodes and the subject's skin below 10 kΩ, a conductive gel was applied to the ground (FPz) and to the 13 electrodes placed in positions F7, Fz, F3, F4, F8, T7, C3, Cz, C4, T8, P3, P4, Pz. Two EMG electrodes were placed on the user's hands to check for the presence of involuntary movements. The following MI tasks were tested:

— MI of right hand (RH): close and open right hand, simulating the clamping movement.

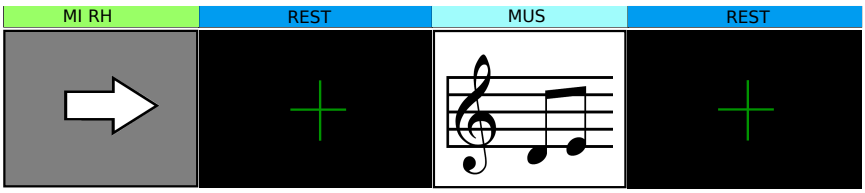


FIGURE 4.7 – Experimental paradigm applied in the investigation phase. We tested different time intervals of 5 s of tasks and 10 s of rest, 3 s/5 s, 3 s/10 s in order to detect the time interval that elicited prominent brain patterns.

— MI of left hand (LH): close and open left hand, simulating the clamping movement.
— Language (LAN): generating words that begins with a specific letter.
— Auditory imagination (MUS): imagining singing a song.
— MI of both feet: move both feet.
— Calculus: imagination of incrementally summing numbers.
— No control (NC): relax.

The tasks were combined in different experimental paradigms, that were tried randomly during the first three sessions (S01, S02 and S03). The subject had to perform the mental tasks following the experimental paradigm that generally consisted in the combination of one or two control tasks interleaved by a no control task. In the no control task (NC) the user was not engaging in any MI task, but she had to achieve a relaxed state focusing her concentration on a cross in the screen. An example of an investigation paradigm is detailed in Figure 4.7. We tried different intervals between tasks in order to identify the interval combination that created more prominent brain responses. The paradigm was repeated 10 times in each run.

Then, to detect the brain pattern of each task, the power spectrum was computed in order to identify the event-related (de)synchronization (ERD/ERS) analysis, introduced in Chapter 1. To create an efficient and adaptive BCI system, the selection criteria of the four tasks are: the most distinguishable on the EEG and the easiest to realize for our pilot. For instance, some tasks such as the calculus create a lot of stress for the subject, the MI of the feet and of the left hand was also really complicated for her and for these reasons were considered as unsuitable tasks.

Finally, at the end of this investigation phase the mental tasks suitable for our pilot were RH, LAN, MUS and NC. These tasks provided a specific brain

pattern in the pilot's EEG, as we can see in the ERD/ERS maps in Figure 4.8, and the subject was comfortable with these tasks.

4.2.2 Training phase

The objective of the training phase was to train the subject to perform the mental tasks selected in the investigation phase. Indeed, in this phase, the pilot had to perform many MI tasks without any feedback, the only aim was to improve her capability to manage the tasks in order to create the training set to calibrate the BCI-game.

The data were collected with the same hardware described in the previous investigation phase (see Figure 4.9), the session took place once a week from the middle of June 2019 to end of August 2019 for a total of 8 sessions. At the beginning (Session S04 and S05) the experimental protocol consisted in 5 runs with the combination of 4 commands, but the subject reported that it was hard because it required a lot of concentration.

Therefore, from Session S06 to S07, the protocol consisted in 4 runs (RH-NC, RH-MUS-NC, RH-LAN-NC and RH-MUS-LAN-NC) and we collected 10 trials per task and run. After these session we tried to reintroduced the task LH. Indeed, the subject at this moment improved her control on the task RH and we wanted to test if this improvement provided also an improvement on the LH task. Hence, from Session S08 to Session S13 the protocol consisted in 5 consecutive runs (RH-MUS-NC, RH-LA-NC, RH-MUS-LA-NC, RH-LH-NC and RH-LH-LAN-NC). The objective was to find the 4-class combination with the highest performance. An illustration of the 4-class experimental paradigm is exemplified in the Figure 4.10, where the control task is represented by a small image in the same view of the game, which is shown in the moment the subject would have to perform it, in order to get the subject used to perform the right task in the game.

In order to detect the ERD and ERS in the EEG associated to the individual mental tasks the EEG signal was bandpass filtered in 6 different frequency bands (8–12 Hz, 16–20 Hz, 20–24 Hz, 28–32 Hz, 32–36 Hz, 36–40 Hz). The ERD/ERS are shown to be active from around 0.5 s after the stimulus and last between 2 and 3 seconds. Therefore we considered epochs of 2.5 s from the mental imagery onset, with steps of 0.5 s, in order to build a BCI system

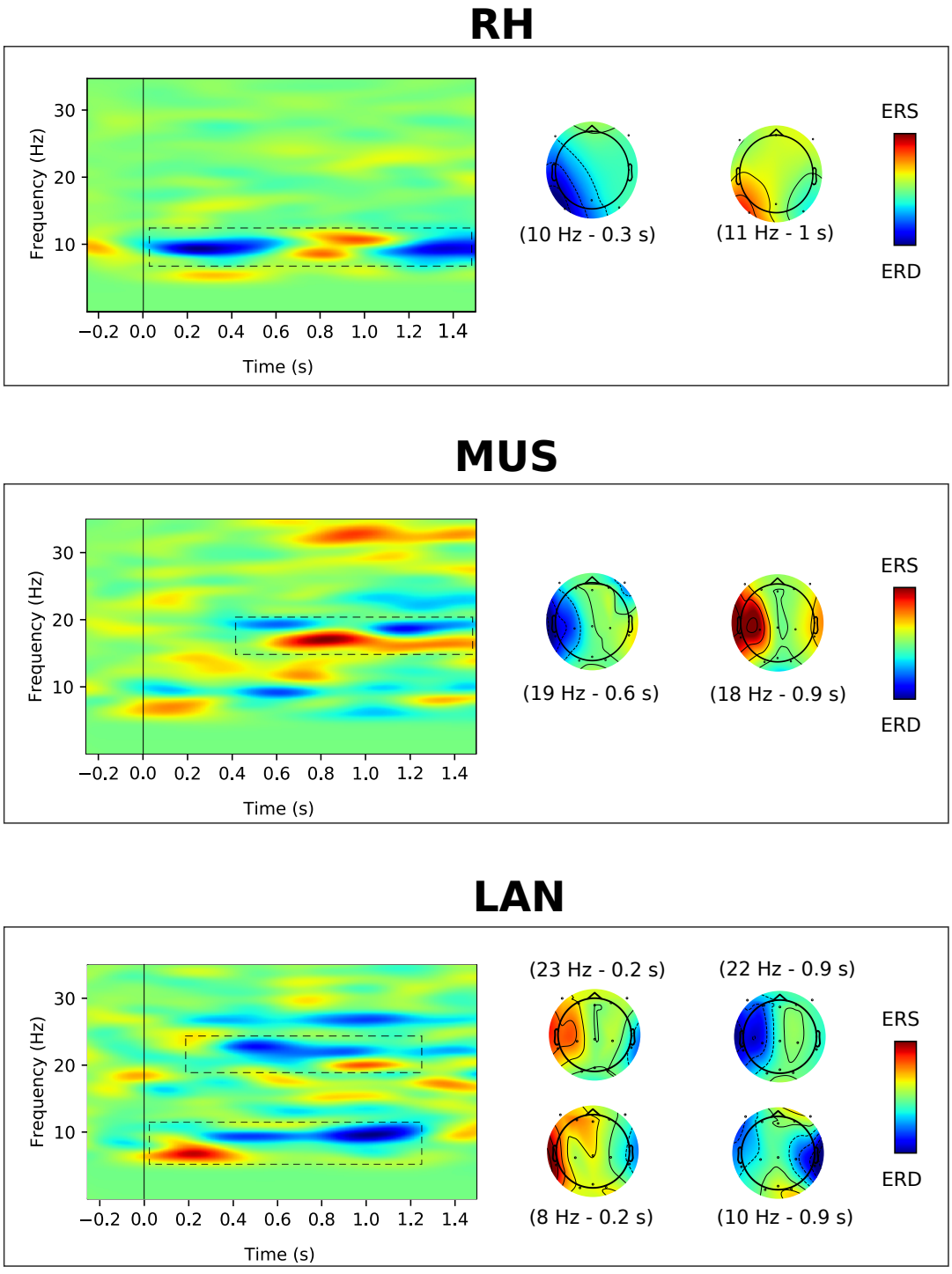


FIGURE 4.8 – ERD/ERS maps calculated for MI of right hand (RH), auditory imagination (MUS) and word association (LAN). For each task the pattern of activation is recognizable by dashed boxes in the frequency-time plot and the scalp topographies indicate the distributions of ERD/ERS at a specific time and frequency.

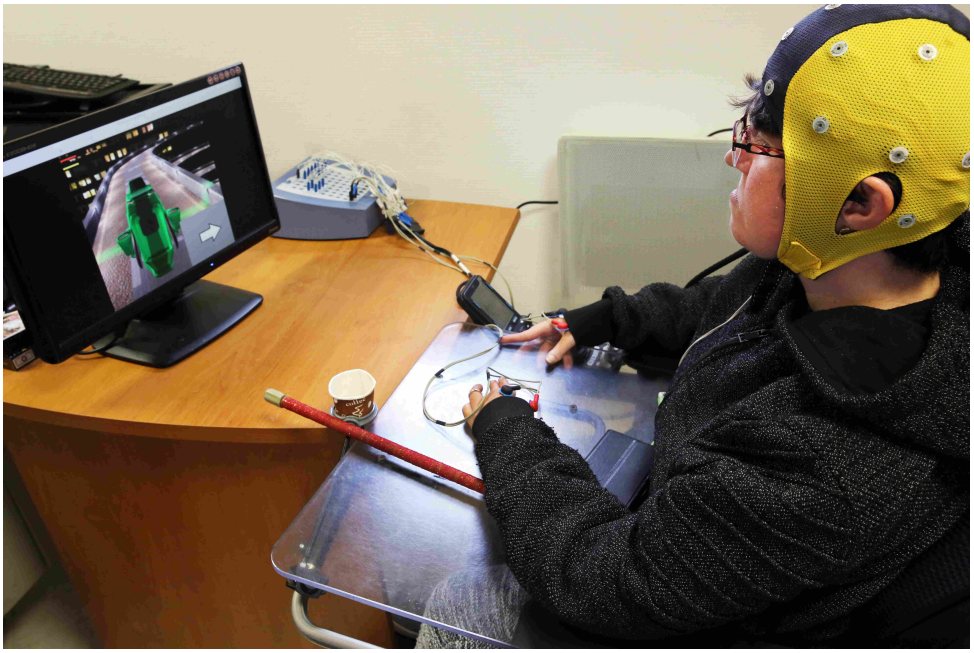


FIGURE 4.9 – Experimental setup of the MI-BCI system during the training phase. The pilot is wearing the EEG cap and the EMG electrodes are placed on her hands.

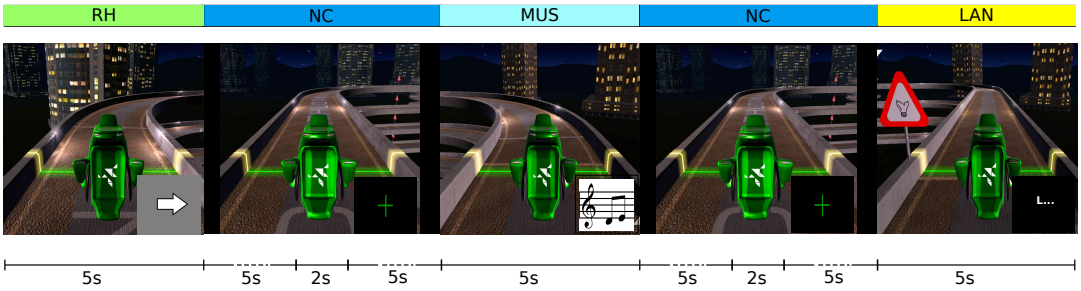


FIGURE 4.10 – Experimental 4-class paradigm applied in the training phase. The user had to perform a control tasks (RH, MUS, LAN) for 5 s. Each task was associated to an image that combines the task image and the corresponding command on the game. The rest interval, corresponded to the no control task (NC), has a total duration of 12 s, after 5 s of the rest interval a green cross appeared on the screen for 2 s to improve the concentration of the pilot.

that reacts as fast as possible to the pilot's intent during the online game. Then a LDA was applied to classify the different tasks, considering the 70 % of the band-power features as training set.

A multi-class confusion matrix was computed to assess the performance reached by each task for the different experimental paradigms. In particular, to analyze the performance of the individual tasks per session, we considered the F-score, that is a statistical measure to evaluate the test's accuracy considering both the precision and the recall (4.3). The *precision* is the number of True Positive (TP) divided by all positive predictions (TP+FP) returned by the classifier, and *recall* is the number of True Positive (TP) divided by the number of all samples that should have been identified as positive (TP+FN). The F-score is more suitable for multi-class problems than the overall accuracy because it is not dependent on True Negative (TN), that can overestimate the performance of the system [Sokolova and Lapalme, 2009].

$$F\text{-}score = 2 \cdot \frac{precision \cdot recall}{precision + recall} \tag{4.3}$$

Figure 4.11 shows the F-score achieved by the individual tasks across session. We can notice that the performance reached in runs with 3 classes is better than the performance with the combination with 4 classes. This trend perfectly reflects the difficulty of the subject to perform runs with 4 tasks, as she declared. For this reason we designed this type of protocol, in which the subject could train on the different tasks gradually achieving 4-task training, that required more concentration and effort.

Generally, we can notice that, independently on the experimental paradigm, the user can manage better the NC and MUS tasks than the RH and LAN. Indeed the NC tasks reached across sessions and paradigms a F-score value never lower than 0.8. Analyzing the results reached for the 4-class combination RH-MUS-LAN-NC we can notice an average improvement of performance from Session S08 to Session S11.

Although there is not evident improvement of the subject performing the tasks, we can notice a gradual increase of the performance across task paradigms, except for specific sessions in which some tasks did not reach a high F-score, as we can see in Figure 4.11.

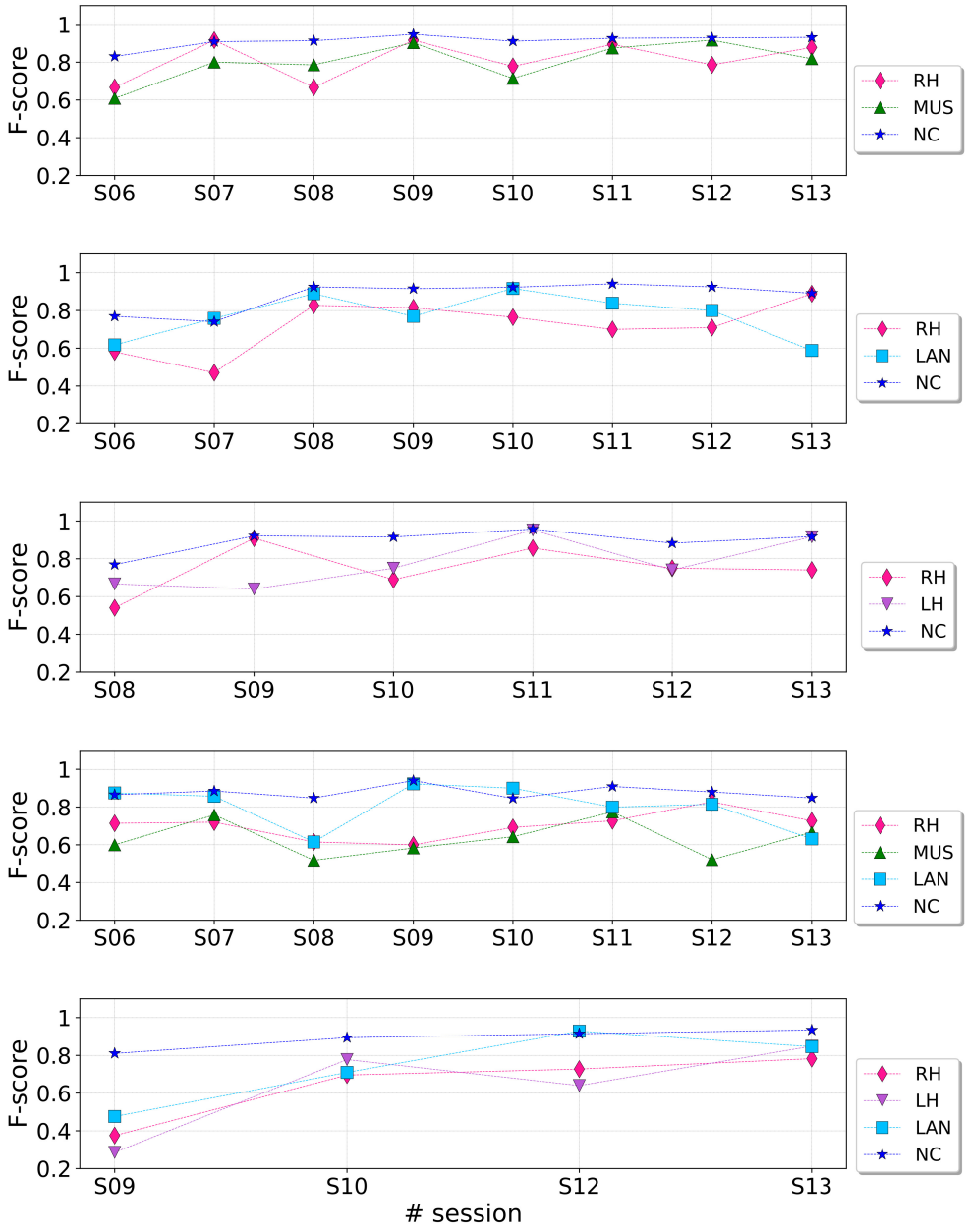


FIGURE 4.11 – F-score values reached by each training paradigm across sessions.

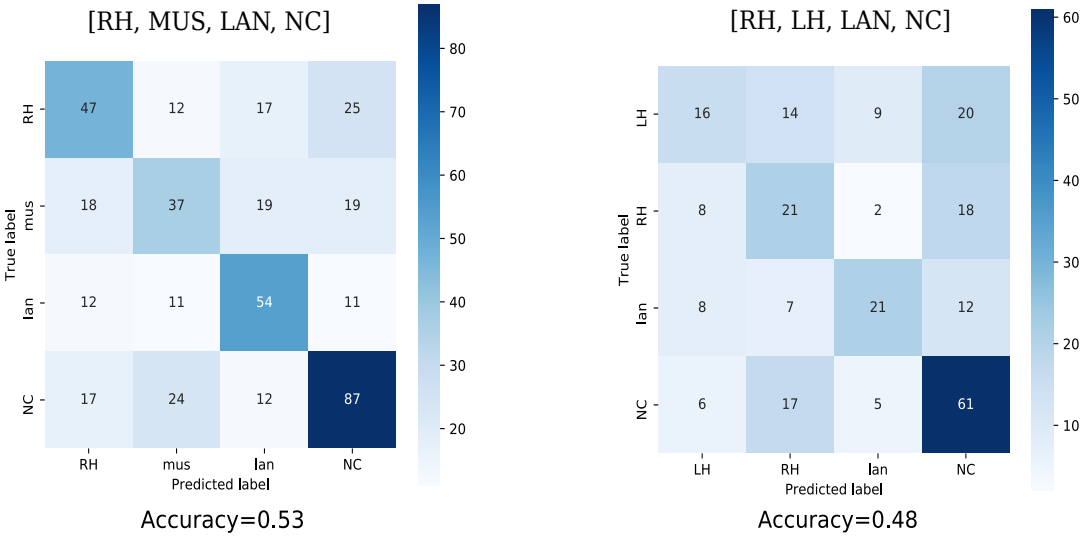


FIGURE 4.12 – Confusion matrix of the two 4-class paradigms tested during the training phase, from Session S08 to Session S13. Each confusion matrix reports the absolute values and relative percentages to evaluate the performance of the LDA classifier. All values on the diagonal represent the correctly classified trials. At the bottom the overall classification accuracy.

Finally, to evaluate the 4-class combination, the confusion matrix across sessions S08 to S13 were computed and shown in Figure 4.12. They show that the results reached with both 4-class combinations are really close. We can observe that with the RH-MUS-LAN-NC paradigm the FN and FP for each class are lower than for RH-LH-LAN-NC paradigm. The accuracy was reported to compare the classification among the 4-class combinations. The accuracy is computed as the sum of the correctly identified classes (TP+TN) over the all the classified classes (TP+TN+FP+FN). The results show a difference of 5 %, with an accuracy value equal to 53 % for RH-MUS-LAN-NC paradigm and 48 % for RH-LH-LAN-NC paradigm.

Therefore, the RH-MUS-LAN-NC paradigm reached the highest performance and was selected as the paradigm to apply in our closed-loop gaming BCI. Finally, the user agreed on this choice because she declared to be much more comfortable with the RH-MUS-LAN-NC combination than with the RH-LH-LAN-NC combination.

The closed-loop BCI game

Figure 4.13 shows an illustration of the closed-loop BCI game. Basically, the EEG signal is acquired from 13 channels and it is bandpass filtered. In parallel the EMG signal is processed in order to detect possible artifacts. Then, each retained epoch is processed to detect the EEG artifact. The EMG and EEG artifact rejection framework will be detailed in the following paragraph. Each processed epoch is a feature vector that will be classified by a LDA classifier, trained with the feature vectors obtained during the training phase. Finally, the classification outputs are mapped to the video-game commands.

However, we encountered some problems to establish the communication between our software developed in Openvibe and the game. The Openvibe software crashed the system, caused by some compatibility and memory issues on the Windows operating system, therefore it did not allow to control the game. As a consequence, we were obliged to use two different computers with two different operating systems: the Linux distribution for the EEG signal processing and Windows for the EEG signal acquisition and game control. Then we established the communication between our processing software and the game by a network connection, allowing us to control the game and avoiding latency or memory problems.

Because of these technical issues it was not possible to test the system before our participation to the Graz BCI race, but we were only able to test it just minutes before the competition.

Artifact rejection scheme

To follow the Cybathlon BCI race regulations, we deployed an artifact rejection framework into the BCI system, that includes both electromyogram (EMG) artifact rejection and electrooculogram (EOG) artifact rejection. The artifact rejection scheme detected the EOG and/or EMG artifacts on the signals and avoided the BCI system to send any control command to the pilot's virtual car for a predefined time interval.

For the EMG artifact rejection two adhesive electrodes for surface EMG were placed on both pilot's hands between the thumb and index fingers. We adopted this configuration because the only motor tasks achievable by the

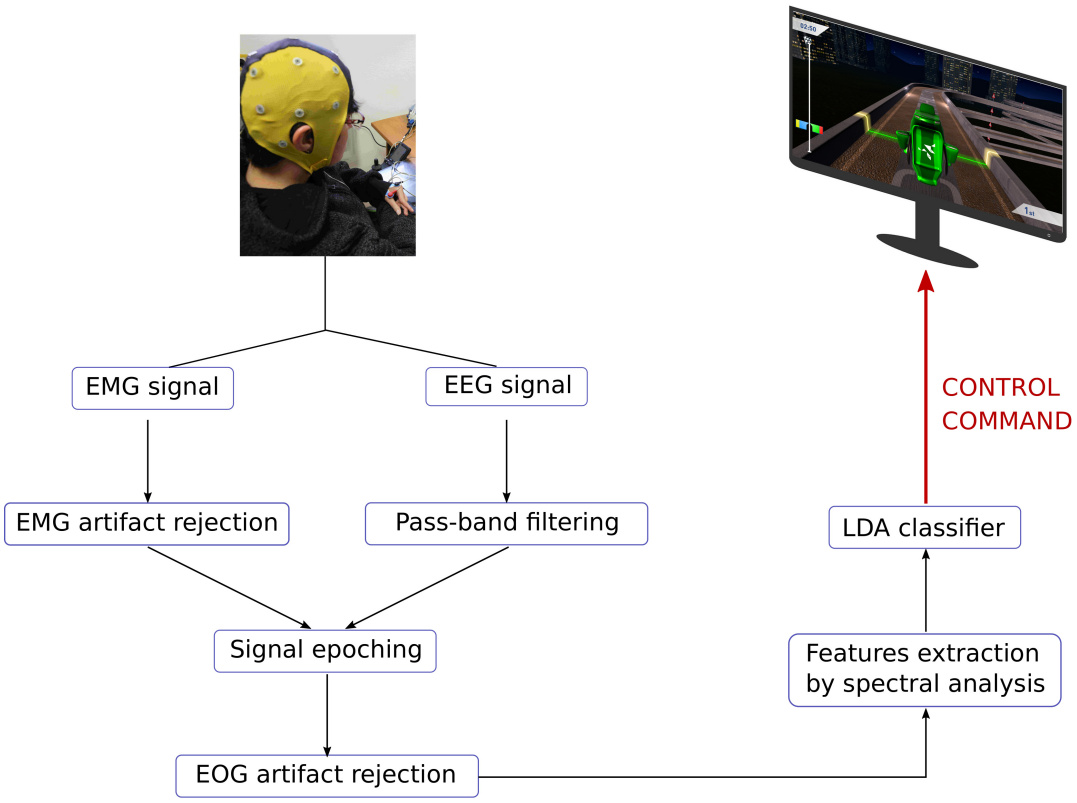


FIGURE 4.13 – Outline of the close-loop BCI system.

user were to close and open the hand, as in a clamping movement. These motor tasks were the real movements that correspond to the RH and LH tasks.

The subject did two acquisitions, in which she performed voluntary hand movements at regular intervals in order to define a release threshold T_{EMG}. Thus, a rejection algorithm was defined, in which for each epoch the average EMG signal amplitude was computed and if it exceed by two standard deviations the threshold T_{EMG}, an EMG artifact was detected, so the epoch was rejected and no command can be sent to the game.

For the EOG artifact rejection the objective was to detect the eye blinking on the EEG signals. In order not to overload the pilot with sensors, the EOG artifact rejection was performed detecting the presence of eye blinking on the frontal EEG electrodes F3 and F4, where the electrode F3 was associated with the left eye and the electrode F4 was associated with the right eye. The artifact rejection was performed processing the EEG signals in the 8–12 Hz frequency band, in which the eye blinks of our subject was more prominent. All samples of individual epochs were labeled following this specific criterion: if

the amplitude exceeded the average amplitude by three standard deviations, the sample was labeled as "False", otherwise was labeled as "True". Then a Boolean sum was applied between the epoch samples of channels F3 and F4. Finally, only the samples labeled as "True", after the Boolean sum, were retained in the processing because not corrupted by eye blinking.

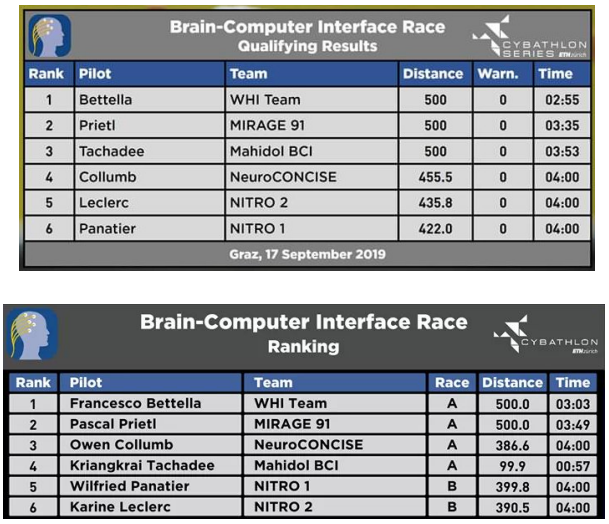
4.2.3 Cybathlon BCI series

The Cybathlon BCI series event took place in Graz in September 2019. This BCI race offered the opportunity to showcase our research and development and gave the pilot an experience of a competition, in preparation for the Cybathlon 2020 event. Six international teams participated to this event and all teams, except the NITRO 1 team and our team (NITRO 2), had previously participated to Cybathlon 2016. Each pilot had to mentally drive a virtual car in a race by his/her mental tasks.

The criterion for winning the game was to complete the track in the shortest possible time, not exceeding 4 minutes. The competition consisted in two phases: Qualifiers and Finals. Two qualifier races of three pilots were organized, the four first pilots in the qualification ranking took part to the final race A, the last two to the final race B.

The official results of the Cybathlon BCI series are shown in Figure 4.14. One pilot was disqualified during the final race. The pilots who reached the first and the second place finished the whole track with a very good timing. We reached the fifth position during the qualifier race and the last position during the final race.

Nevertheless, the Graz experience was really useful for the future improvement of the system. We had the possibility to test our system in real life conditions, we understood the limits of our system and on what we need to work on to become more competitive for the Cybathlon race.



Brain-Computer Interface Race
Qualifying Results

Rank	Pilot	Team	Distance	Warn.	Time
1	Bettella	WHI Team	500	0	02:55
2	Prietl	MIRAGE 91	500	0	03:35
3	Tachadee	Mahidol BCI	500	0	03:53
4	Collumb	NeuroCONCISE	455.5	0	04:00
5	Leclerc	NITRO 2	435.8	0	04:00
6	Panatier	NITRO 1	422.0	0	04:00

Graz, 17 September 2019

Brain-Computer Interface Race
Ranking

Rank	Pilot	Team	Race	Distance	Time
1	Francesco Bettella	WHI Team	A	500.0	03:03
2	Pascal Prietl	MIRAGE 91	A	500.0	03:49
3	Owen Collumb	NeuroCONCISE	A	386.6	04:00
4	Kriangkrai Tachadee	Mahidol BCI	A	99.9	00:57
5	Wilfried Panatier	NITRO 1	B	399.8	04:00
6	Karine Leclerc	NITRO 2	B	390.5	04:00

FIGURE 4.14 – Cybathlon BCI series ranking.

4.2.4 Discussion and perspectives

The long training phase and the BCI series in Graz provided us an enriching experience to understand the limits that these types of system present, the factors that influence the performance and how the performance can be improved.

We would like to underline that we had only three months to design the system, to adapt it as best as possible to the pilot and to train her, which is not a very long period for the preparation to this type of competition.

We did a long phase of training but we could only train her to play the game itself for a few sessions (2 or 3), before the competition. For sure, learning to use the system and learning to "play" are two different tasks and therefore imply different levels of concentration. For instance, the day of the competition, we noticed that during the last minutes, the pilot had more difficulties to stay concentrated. Concentration skills during the game could have been improved if we had had more time to train the pilot with the game.

There were many factors that influenced the stress condition of the pilot, impacting her ability to concentrate and consequently her performance. For instance, during both investigation and training phases, the acquisition took place in a standard room in a living center, as mentioned before. The room was not a room equipped for EEG experiments and not shielded for external

sounds, consequently many times the acquisition were corrupted by external sounds that caused many distractions for the subject. Moreover all the training phase took place in summer and therefore in very hot conditions, this condition decreased the pilot's abilities of concentration mainly caused by discomfort due to wearing an EEG cap with gel during an extremely hot period.

Also, during the BCI series in Graz, many factors influenced her stress, such as competition stress, travel and others logistic problems. Indeed, this experience highlighted the problems faced by disabled people, particularly in terms of logistics (adapted transport and infrastructure) and special needs (lifts, wheelchairs, etc.). It also led us to realize the stakes of the organization of such an event.

Another factor that probably induced an increase of the overall stress of the pilot was the presence of many people and of noise during the competition. Indeed the competition took place in an amphitheater room and each pilot was positioned in front of the public, and during the competition a person commented the race.

In these BCI series, we noticed a great deal of variability between pilots. For example, residual motor abilities were highly variable from one pilot to another. Some drivers could fully use their arms, others could not move at all. Another reflection is that disabilities could be congenital or acquired. The challenge is to create a system that can be adapted to all situations and this confirms the importance of personalizing BCI systems, to tackle the needs of each user in any situation, such as a competition or real life.

The participation to this competition was a really exciting challenge and provided us a very informative experience in the development of a BCI for a disabled person. We understand that the role of the user is fundamental in a SMR-BCI system, confirming the need to develop user-centered systems in particular for disabled people that present different needs, based also on their disabilities.

After the BCI series there were many aspects that we would have liked to improve in our system. Firstly, train the pilot to play the game with external disturbance (noise and people) and improve her concentration capability, in order to maintain her maximum concentration until the end of the track. Moreover it would have also been productive to allow the pilots to train

against each other in order to simulate real competitions. Secondly, change the hardware system in order to have a more practical equipment, avoiding the need to establish a network connection between the signal processing software and the game.

We started to work on these aspects but we had to stop the training with our pilot because of the Covid-19 pandemic and, sadly, the Cybathlon competition in May 2020 was postponed and it is still not clear whether it will take place or not.

4.3 Conclusions

In this chapter we presented two user-centered BCI implementations. The first concerned an adaptive c-VEP BCI speller tested on healthy people. The second concerned a MI-BCI system for an impaired user, in the contest of a BCI competition.

These works aimed to underline the need of an user-centered system providing an experience on how this challenge can be tackled in different types of systems, either a speller or an SMR-BCI, with different types of user.

In the case of the adaptive c-VEP speller the objective was to develop a method in order to investigate the influence of some stimulation parameters on the performance of a dataset of nine healthy subjects. Hence, the work was focused on the research and development of an algorithm and software addressed to define the most adapted stimulation parameters for each subject and for each session.

On the other hand, for the development of a MI-BCI it was fundamental to define the mental tasks more suitable for the user. Therefore the work was focused on the definition of a specific training protocol in order to improve the ability of the subject to manage her mental tasks.

5 Auto-Calibration of c-VEP BCI by Word Prediction

In Chapter 2 we discussed the limitations caused by the calibration phase in VEP BCI spellers and we provided an analysis of methods known in literature to avoid this expensive phase. Then, in Chapter 4 we proposed to replace the standard calibration phase with an adaptive parameter setting phase in a c-VEP BCI speller, in order to open the way for the development of an user-centered system.

In this chapter we present an auto-calibration c-VEP system, in which the fundamental properties that characterize the VEP response are exploited to predict the full word using a dictionary, eliminating the traditional calibration phase.

5.1 The auto-calibration method

To face the limits of BCI spellers caused by long and tedious calibration phases, we propose an auto-calibration method for a c-VEP BCI system, that allows to spell words and uses word prediction to avoid the traditional calibration phase. The framework of the method is shown in Figure 5.1. The detailed explanation of the auto-calibration method (AC) is as follows.

1. The average response of the first character $\boldsymbol{X_{a1}} \in \mathbb{R}^{c \times s}$ is computed over N stimulus cycles, where c is the number of channels and s the number of samples of the stimulus cycles. $\boldsymbol{X_{a1}}$ is computed using the time-windowed EEG recorded while the user is gazing at the first character, but since the system has not been calibrated, the first character is unknown at this step. For the first character the relative position is 0.

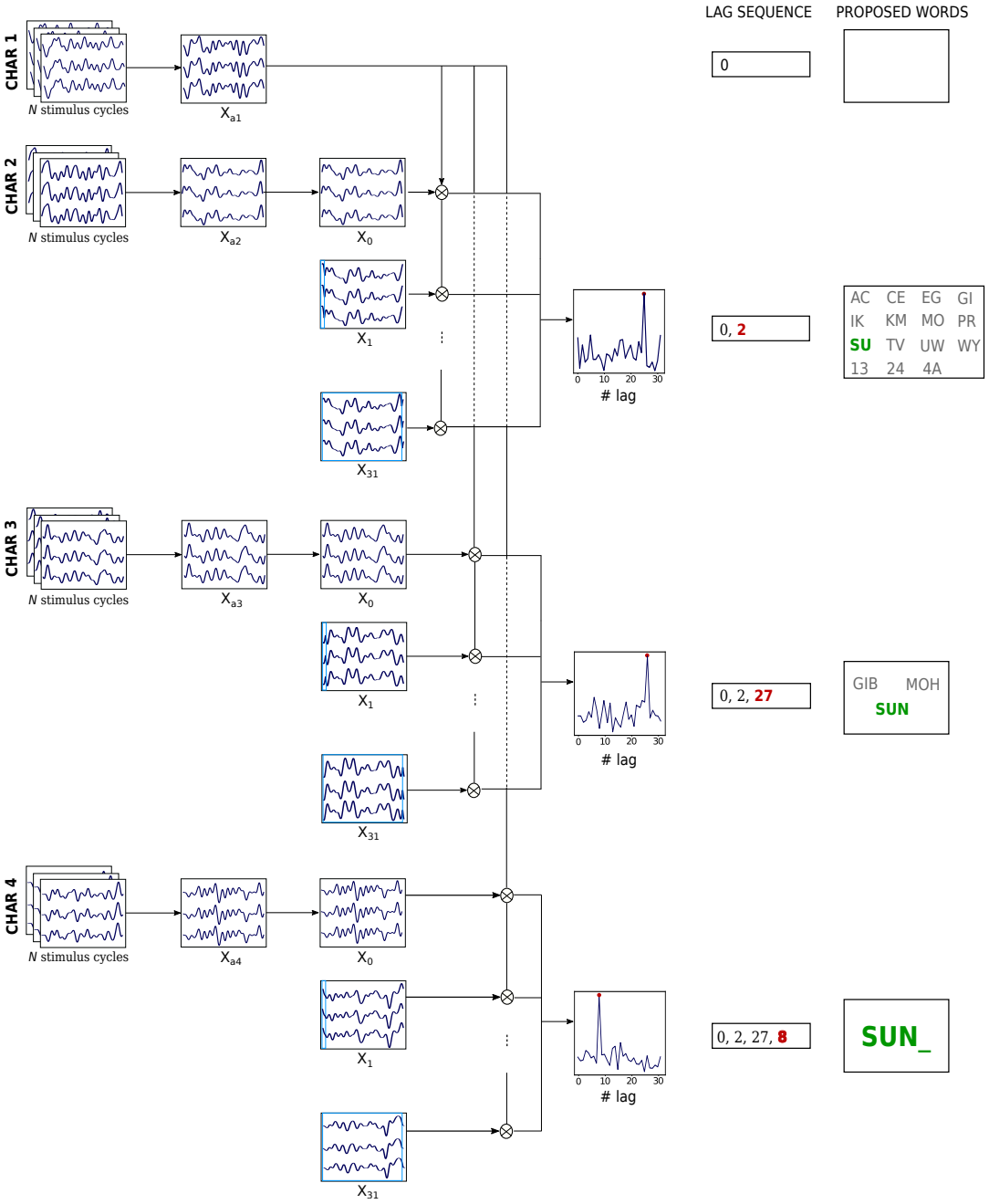


FIGURE 5.1 – Outline of the auto-calibration method for a c-VEP BCI to spell the target word "SUN_".

2. For the second character, the average response $\boldsymbol{X_{a2}} \in \mathbb{R}^{c \times s}$ is computed over N stimulus cycles. $\boldsymbol{X_{a2}}$ is shifted by $(l \cdot \tau)$ time samples, where τ is the time lag between two consecutive characters and l is the index of the corresponding character in the virtual keyboard. This produces L shifted averages $\boldsymbol{X_l}$ (5.1), $l = 0, ..., L-1$, where L is the number of characters on the virtual keyboard.

$$\boldsymbol{X_l} = \boldsymbol{X_{a2}}(t - (l \cdot \tau)) \tag{5.1}$$

3. The correlation ρ_l between the initial average response $\boldsymbol{X_{a1}}$ and the average response of the second character $\boldsymbol{X_l}$ is computed.
4. Using the lag l_{max} (5.2) which produces the maximum correlation between $\boldsymbol{X_{a2}}$ and $\boldsymbol{X_{a1}}$, the relative position of the second character with respect to the first is computed.

$$l_{max} = \arg\max_{l} corr(\boldsymbol{X_{a1}}, \boldsymbol{X_l}) \tag{5.2}$$

5. Among all pairs of characters separated by l only retain those corresponding to the beginning of valid words within a dictionary.

This method is repeated for the following characters, until we are left with a single word. At that moment, we will have recovered the original letter, and the absolute position of $\boldsymbol{X_{a1}}$ can thereafter be used during the computation of the time lag.

In order to guarantee the uniqueness of each word belonging to the dictionary a character "_" is added at the end of each word, in this way each word has a unique sequence of lags, independently of the type and size of the dictionary. Moreover it allows the system to understand when the spelling of a word is concluded.

The auto-calibration method can be improved considering not only the most correlated character, as explain at the step (iv), but also the second most correlated one. However the latter is taken into account only if its correlation value is higher than a specific percentage P_c of the correlation value of the most correlated character. The second most correlated character gives the lag $l_{2,max}$, as shown in Figure 5.2.

The coefficient $\varphi_{n,k}$ is introduced, for each character of the word excluding the first ($n = 2, ..., N_{chars}$) and for the k^{th} most correlated character.

$$\varphi_{n,k} = \frac{corr(\boldsymbol{X_{a1}}, \boldsymbol{X_{l_{k,max}}})}{corr(\boldsymbol{X_{a1}}, \boldsymbol{X_{l_{max}}})} \tag{5.3}$$

Clearly the coefficient $\varphi_{n,1}$ is always equal to 1, because it corresponds to the character with highest correlation ($l_{max} = l_{1,max}$). Then, we can use the coefficient $\varphi_{n,2}$ of the character corresponding to $l_{2,max}$ as a representation of the distance of its correlation with respect to the highest one, and we can apply the threshold P_c to choose whether it has to be considered as a spelling candidate or not. The system will give as output the word belonging to the dictionary with the highest value of φ_{path_i} (5.4), computed as the average of the $\varphi_{n,1}$ and $\varphi_{n,2}$ in the path i through the $(N_{chars} - 1)$ characters. Each path i is obtained combining all the possible suggested lags (l_{max} and $l_{2,max}$).

$$\varphi_{path_i} = \frac{\sum\limits_{n \in path_i} \varphi_{n,1} + \sum\limits_{n \in path_i} \varphi_{n,2}}{N_{chars} - 1} \tag{5.4}$$

5.2 Data and experiments

We investigated data from the public dataset of [Spüler et al., 2012]. Nine healthy subjects participated in the c-VEP BCI experiment. Each subject took part in two identical sessions. For five subjects the session was performed in the same day, for the other four on different days. For a complete subject review refer to [Spüler et al., 2012].

The EEG signal of the subjects was recorded from a Brainproducts Acticap system with 32 channels, with a g.tec g.USBamp (600 Hz sampling rate). They used 30 electrodes located at Fz, T7, C3, Cz, C4, T8, CP3, CPz, CP4, P5, P3, P1, Pz, P2, P4, P6, PO9, PO7, PO3, POz, PO4, PO8, PO10, O1, POO1, POO2, O2, OI1h, OI2h and Iz. The remaining two electrodes were used for electrooculography (EOG). The reference electrode was positioned at Oz and the ground electrode was positioned at FCz. Each session consisted of 10 stimulus cycles per target. In each run the subject had to spell 32 targets on a virtual keyboard twice, using the c-VEP BCI system, so in total

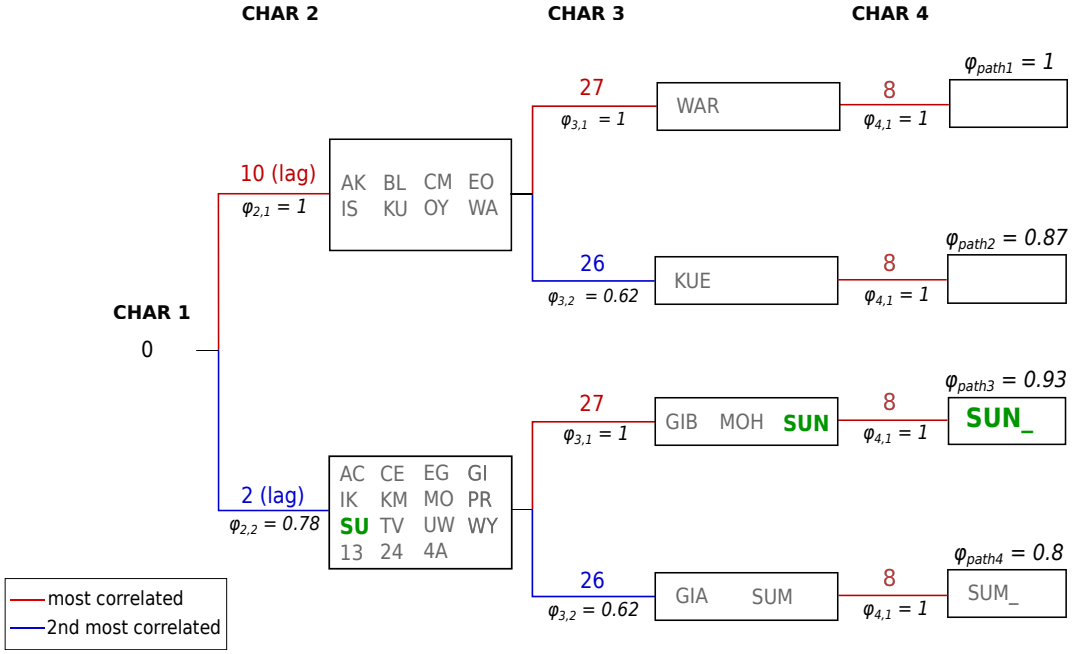


FIGURE 5.2 – Spelling of the word "SUN_" applying the improved auto-calibration method (see Section 5.1). In red the lag l_{max} of the most correlated character and in blue the lag $l_{2,max}$ of the second most correlated character, with their corresponding values of $\varphi_{n,1}$ and $\varphi_{n,2}$, where n indicates the character position in the word. The figure shows that for characters 2 and 3 two lags are found, while for character 4 only the most correlated lag is considered. This is because the values of $\varphi_{2,2}$ and $\varphi_{3,2}$ are higher than the threshold P_c, while $\varphi_{4,2}$ is lower. At the end, the system proposes the word belonging to the dictionary with the highest value of φ_{path}, that respects the lag sequences defined during the spelling.

one session consisted in 640 trials. The arrangement of the targets in the keyboard was conformed to the principle of equivalent neighbors [Bin et al., 2011]. Each target was modulated by a 63-bit binary m-sequence with a stimulus presentation rate of 60 Hz, hence the length of a stimulus cycle was $t_s = 63/60 = 1.05$ s. The modulation sequence of each target was shifted by 2 bits with respect to its preceding target, hence the time lag between two consecutive targets was $\tau = 2/60 = 0.033$ s.

5.2.1 Offline experiments

The signals were preprocessed using a Butterworth filter between 2 and 15 Hz. We simulated the spelling of four different groups of 5 words called *3-char*, *4-char*, *5-char* and *sentence*. Each group has a different number of characters per word and consequently a different total number of characters. The composition of the groups of words is detailed in Table 5.1. The words belonging to *3-char*, *4-char* and *5-char* groups are randomly selected from a free corpus of ten thousand English words downloaded from the Wikipedia corpus, that was used as dictionary for the auto-calibration method. For the *sentence* group, the sentence chosen to be spelled was "BRAIN_COMPUTER_INTERFACE_LETS_EVERYONE_COMMUNICATE". The dictionary was modified appending the character "_"at the end of each word, as explained in Section 5.1.

The channels OI1h, OI2h and Iz are not considered for our experiments because the amplitude of the EEG signal is too low. We conducted experiments with different methods and we use the following convention to name them: standard calibration (C), explained in Section 2.1.1, the auto-calibration method considering only the most correlated character (AC1) and the auto-calibration method considering the two most correlated characters (AC2), explained in Section 5.1.

For the experiments with method C we performed the same pre-processing as applied in [Spüler et al., 2012]. Then we applied the CCA spatial filter to N stimulus cycles of the character "A" to compute an average absolute response X_1, used as reference template. Finally, we followed the process described in Section 2.1.1. Each experiment was parameterized by the number of stimulus cycles, in two different ways. The former experiments evaluated

TABLE 5.1 – Composition of the groups of words applied in the experiments.

Word group	# total chars	# words	# chars per word
3-char	15	5	3
4-char	20	5	4
5-char	25	5	5
sentence	51	6	[5, 12]

the performance of methods C, AC1, AC2 with the same number of stimulus cycles for each group of words. In particular, for the C method the number of stimulus cycles applied in the calibration phase and offline spelling phase is the same. The latter experiments evaluated the performance of AC2 method to spell the same groups of words with a number of stimulus cycles fixed to 10 for the first character and following the parameterized number of stimulus cycles for all the other characters belonging to each word. This can be considered as an option of the AC2 method, called AC2* method.

5.3 Results

5.3.1 Comparison between methods

Figure 5.3 shows the accuracy reached by all subjects of the dataset to spell different groups of words applying the C method, explained in Section 2.1.1, the AC1 and AC2 methods, explained in Section 5.1. The composition of each group of words is described in Table 5.1.

The accuracy is computed as the ratio of characters well spelled parameterized by the number of stimulus cycles. The results reached in Session 1 show that the average accuracy achieved with the C method is higher than the average accuracy reached applying both the AC1 and AC2 methods, except for the *3-char*, *4-char* and *5-char* groups for which the average accuracy with the AC2 method is higher for 3 and 4 stimulus cycles, even if the dispersion is larger with respect to C method.

On the contrary, in Session 2 the average accuracy attained with the C method is lower than the average accuracy reached with the AC2 method for all the tested groups of words and for all the numbers of stimulus cycles.

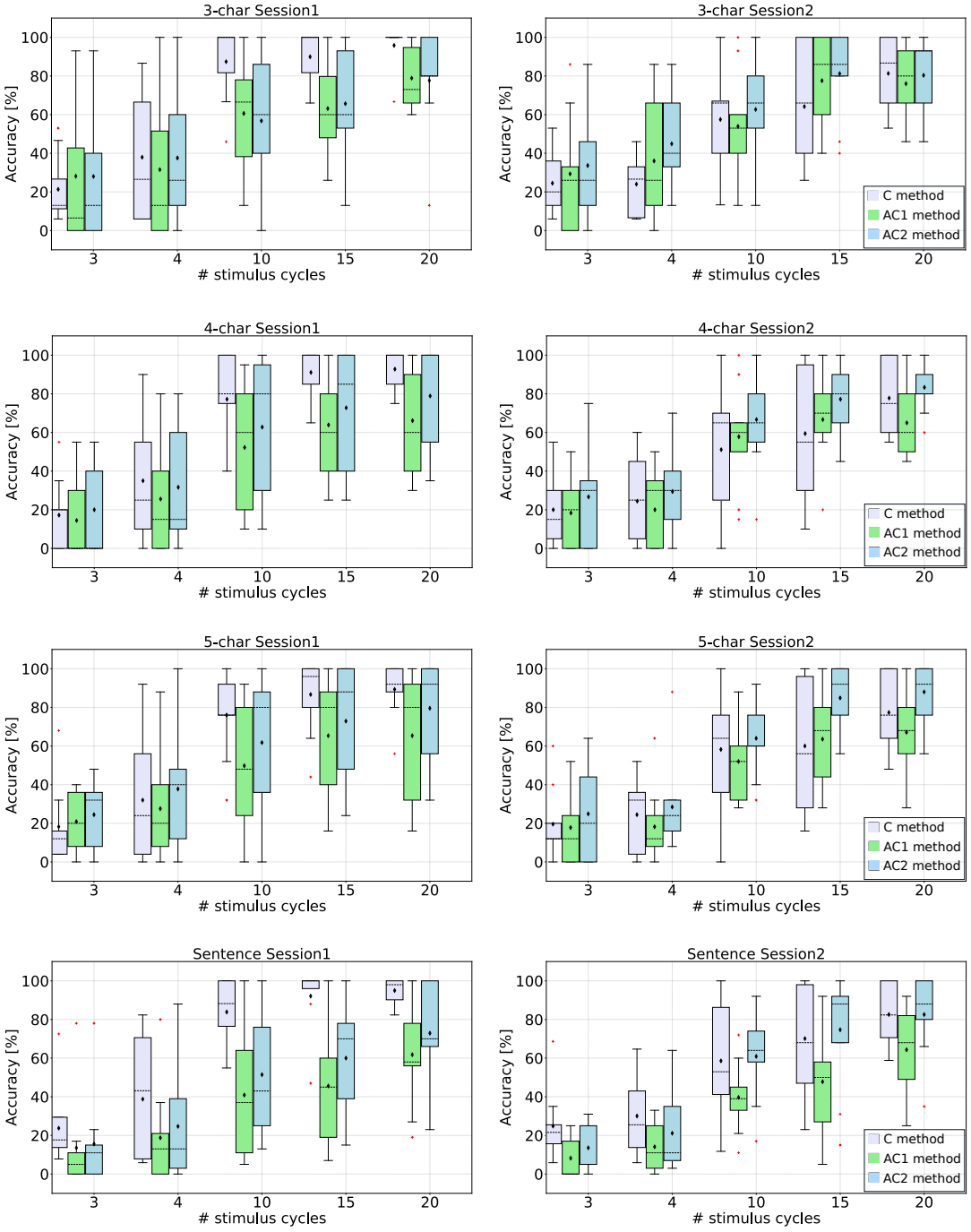


FIGURE 5.3 – Boxplots of accuracy reached by all subjects for spelling four different groups of words parameterized by the number of stimulus cycles in Session 1 and Session 2, applying the standard calibration (C) in violet, the auto-calibration method considering only the most correlated character (AC1) in green, and the auto-calibration method considering the two most correlated characters (AC2) in light blue. The box edges represent quartiles, diamond markers represent the mean and dash lines denote the median. Outliers are marked with red crosses.

The only exception is for the spelling of a sentence with 3 and 4 stimulus cycles, for which the average accuracy of the C method is the highest.

Generally, comparing the results reached across sessions, we can observe that the average accuracy improved increasing the number of stimulus cycles, for all methods, and the dispersion in Session 1 is larger with respect to Session 2, in particular for AC1 and AC2 methods.

Moreover it is evident that the AC2 method performed better than the AC1 method for all the tested cases reported in Figure 5.3.

5.3.2 Method AC2

To evaluate the performance of the AC2 method we selected the subset of six subjects with good performance, reaching an average accuracy higher than 30 % across the number of stimulus cycles and the groups of words. The subjects not included in this subset are AE, AA, AH, for more details refer to [Spüler et al., 2012].

In Figure 5.4 is presented the performance reached with the AC2 method across sessions. The trends in Figure 5.4 confirm the improvement of the performance increasing the number of stimulus cycles, except for the *3-char* and the *5-char* groups, for which the performance decreased in Session 2 from 15 to 20 stimulus cycles. Specifically for the *3-char* group the average accuracy is equal to 93 % with 15 stimulus cycles and equal to 85 % with 20 stimulus cycles, while for the *5-chars* group the corresponding values are respectively 95 % and 91 %. Nevertheless, it is evident that the results in Session 2 are less spread than in Session 1, as we can also observe in Figure 5.3.

In Session 1 the *4-char* and the *5-char* groups obtained better performance on average with respect to other groups. Instead in Session 2 the accuracy reached for the *3-char* group is higher than the other groups, especially for 3 and 4 stimulus cycles. On the contrary we can observe that for more than 4 stimulus cycles the performance for all groups is really close, except for the *4-char* group with 15 and 20 stimulus cycles. Comparing the two plots it is clear that the variability of the performance is larger for low numbers of repetitions (3 and 4 stimulus cycles) across groups and sessions.

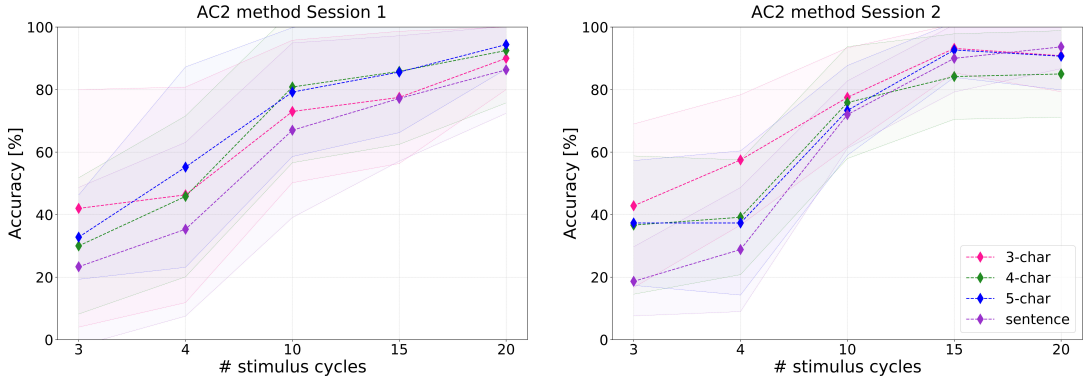


FIGURE 5.4 – Evolution of AC2 method performance over sessions. Diamond markers represent the average accuracy reached across repetitions for a subset of subjects, with the *3-char* group (pink), the *4-char* group (green), the *5-char* group (blue) and the *sentence* group (purple). The transparent zones indicate the standard deviation values. The subset of subjects includes all subjects that reached an average accuracy higher than 30 % across stimulus cycles and groups of words.

In order to evaluate a possible improvement of the performance with a lower number of stimulus cycles and, at the same time, avoiding a very long stimulation time, the methods AC2 and AC2* are compared in the boxplots shown in Figure 5.5. The performance of both methods is compared across sessions on the same subset of best subjects evaluated in Figure 5.4. The average accuracy increases with the AC2* method, reaching values always higher than the ones obtained with the AC2 method, usually close to the accuracy reached with 10 stimulus cycles. This is especially evident for the Session 2, with the exception of the *sentence* group, indicating a particular benefit for the performance with a lower number of stimulus cycles.

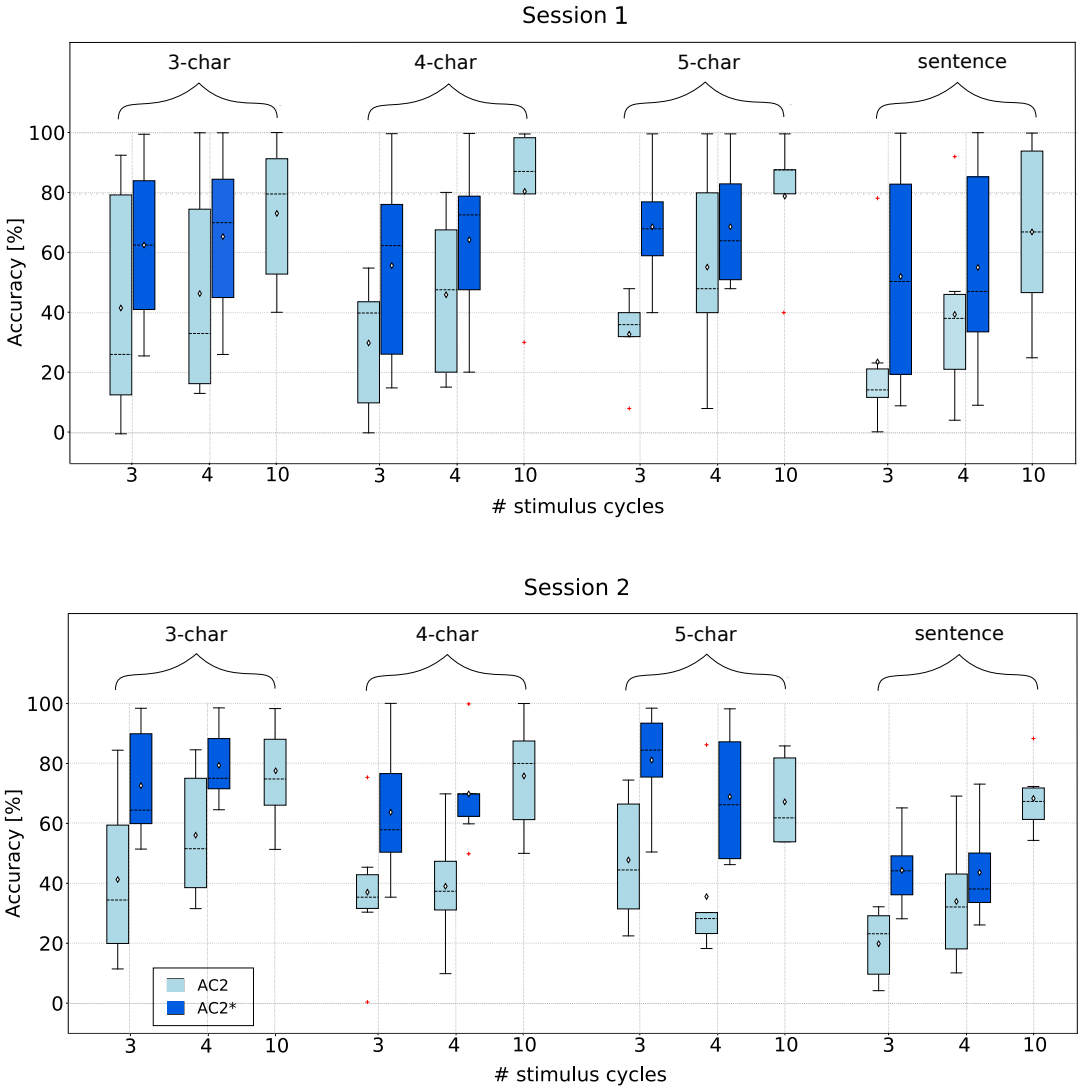


FIGURE 5.5 – Boxplots of accuracy reached with the AC2 method (light blue) and the AC2* method (dark blue) in Session 1 and Session 2. The box edges represent quarterlies, diamond markers represent the mean and dash lines denote the median. Outliers are marked with red crosses.

5.4 Discussion

The objective of our work is to develop a method for a c-VEP BCI with no calibration phase. The AC1 and AC2 methods proposed in our work were tested for four different groups of words to evaluate the stability of the method when varying the number of characters of a single word and the total number of characters to spell.

The results shown in Figure 5.3 and, in particular, Figure 5.4 demonstrate that there is no clear prevalence in terms of number of stimulus cycles with respect to the number of characters in the word. However, it appears that the performance is lower with a low number of stimulus cycles for a high number of characters, but when increasing the number of stimulus cycles, the performance is comparable for all the groups of words, except for *sentence* group. Indeed it is evident that the performance for the *sentence* group is lower with respect to other groups of words. In this case we simulated the spelling of a sentence with a total of 51 characters and 6 words with a maximum of 12 characters per word. We chose this sentence with long words in order to evaluate the efficacy of the method in a more complex scenario. In fact with longer words the risk of not reaching the correct spelling of the target word increases, because the probability of not correctly identifying a character increases and, consequently, the wrong identification of the word in the dictionary. This justifies the lower performance of the spelling of the *sentence* group compared to other groups of words.

This aspect may be improved including, for example, the error correction [Speier et al., 2016] in an online implementation of the method, to reduce the wrong target word identification caused by the erroneous selection of single characters. Furthermore, in online applications, the subject may initially spell shorter words in order to become familiar with the system, improving the ability to use it and consequently improving the repeatability of the VEP responses, thus limiting errors also when spelling long words.

The repeatability of the VEP responses is fundamental for c-VEP BCI systems and in particular to achieve good performance with the proposed auto-calibration method. For instance, in Figure 5.4 we can notice a decrease of the average performance of the *3-char* and the *5-char* groups in Session 2 from 15 to 20 stimulus cycles, caused probably by the lack of repeatability of VEP responses of some subjects for some specific characters. Moreover, Figure 5.3 shows a consistent dispersion with all methods, especially in Session 1, demonstrating that for some subjects it is possible to obtain very good results also with a lower number of stimulus cycles, but for others it is difficult to reach good results. Figure 5.6 shows the repetitions of each character realigned with respect to the unshifted m-sequence for one of the best and one of the worst performers in Session 1, in terms of accuracy (subject AD and subject AE, respectively). We can notice visible bands of activation (red

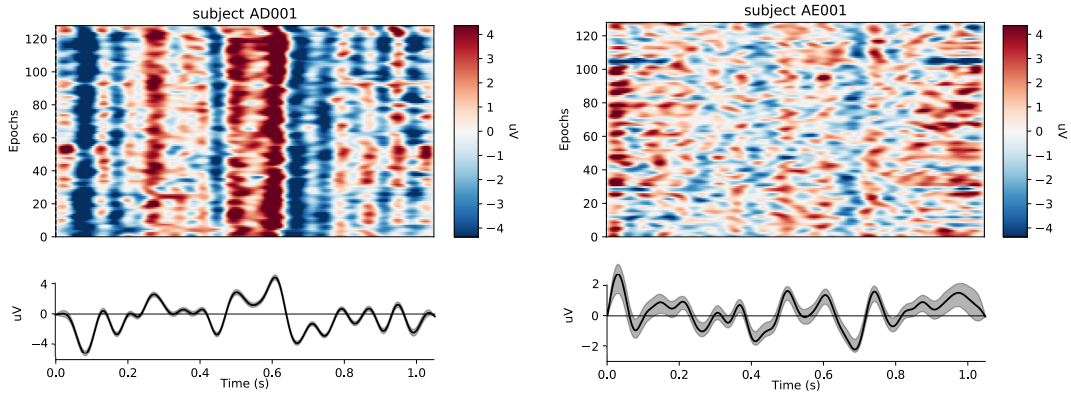


FIGURE 5.6 – The first four stimulus cycles of each target character realigned with respect to the unshifted m-sequence and the average realigned response with the standard deviation in Session 1, for subject AD and subject AE, respectively.

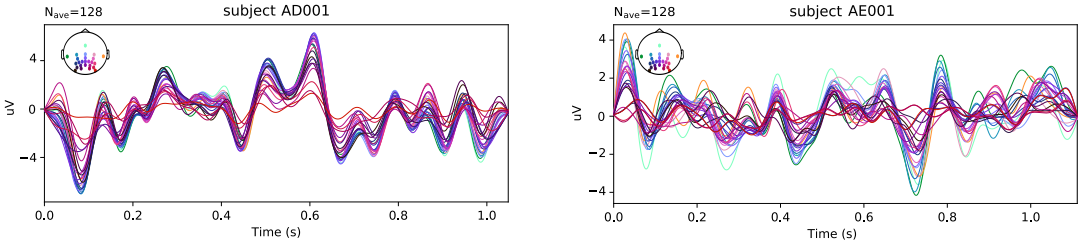


FIGURE 5.7 – Average response of each channel over the first four realigned stimulus cycles of each character in Session 1, for subject AD and subject AE, respectively.

zone) for all the repetitions for the subject AD and an average with a low standard deviation. On the other hand, for subject AE it is not possible to define specific zones of activation because there is a visible variability between the repetitions, also demonstrated by the standard deviation computed over all stimulus cycles, which is greater than for subject AD. Moreover Figure 5.7 shows the average response over all the realigned repetitions for each channel for the same two subjects. We can notice that for subject AE some electrodes in the occipital part did not work well. For sure they influenced the low accuracy reached for the subject in the offline experiments. Instead for subject AD all the electrodes have the same response shape, with a difference in amplitude. So we can affirm that the performance of the auto-calibration method is influenced by the quality of the data. In particular it is important to have a good VEP response repeatability between the stimulus cycles for each character.

The quality of the data is fundamental to build a high performance BCI system. There are many factors that can influence the EEG variability. Some of them can be controlled, for example the design of the system and the hardware used for the acquisition of EEG signals, the application of specific algorithms and the control of some external factors like noise. But we can not control the state of subject, like the loss of concentration, visual fatigue, distraction and motivation [Kleih and Kübler, 2015]. The EEG variability influences the performance of the BCI system from a session to another. With the auto-calibration method it is possible to face the problem of inter-session variability because for each session the system is not calibrated with signals acquired in a different time, but it is based on the data acquired during the specific session considering all the possible factors that can influence it. This explains why with a lower number of stimulus cycles, for many subjects, the accuracy reached with the auto-calibration method is higher than the accuracy reached with the standard calibration method. However, we can notice that the performance with the calibration method increased with the number of stimulus cycles. Indeed, generally, in c-VEP BCI the performance increases with a longer calibration phase [Bin et al., 2011; Wei et al., 2016].

Analyzing the results showed in Figure 5.3 it is evident that with the auto-calibration method the accuracy increased when considering the two most correlated characters. A possible explanation is that if we analyze the stimulus modulation applied to collect the dataset [Spüler et al., 2012], each character flashes according to a 63-bit binary sequence and the sequence generated for each character is 2-bit circularly shifted with respect to the sequence of the consecutive characters [Bin et al., 2011]. This means that the time shift between two consecutive characters is 2 bit/60 Hz = 0.033 s. Therefore the timing precision of the VEP response is fundamental to distinguish the response of each character. In fact it is possible that some stimulation parameter can influence the VEP response, such as stimulus proximity and the lag between two adjacent stimuli [Wei et al., 2016]. The subject could be distracted by the interference caused by consecutive characters in the keyboard which adds the difficulty to focus the attention on the target character. This is also proved by the fact that when applying the AC2 method the couple of lags detected corresponds to two consecutive character, as shown for example in Figure 5.2.

5.4.1 Performance evaluation

To compare the potential effectiveness of this method with the actual state-of-art of online c-VEP BCI, we compute the theoretical information transfer rate (ITR) [Wolpaw et al., 2002], following equation (5.5). The ITR returns the amount of information contained in a selection in bits per minute (bpm).

$$ITR = \left(\log_2 N + P\log_2 P + (1-P)\log_2\left(\frac{1-P}{N-1}\right)\right)\cdot\left(\frac{60}{T}\right) \tag{5.5}$$

Where N is the number of classes, equal to 32, which is the number of characters in the virtual keyboard, P is the accuracy (the percentage of characters correctly spelled) and T is the average time required for the prediction. We computed the theoretical ITR for the spelling of all word groups comparing the values reached with the AC2 method and the AC2* method, across sessions, for the subset of best subjects.

The time T was computed as the average time required to select a character, taking in account the number of stimulus cycles for the total number of characters for each group of words. For instance for the *3-char* group the time T is computed as the number of stimulus cycles multiplied by the duration of one stimulus cycle (1.05 s) for the AC2 method. Instead for the AC2* method we computed an average T, taking in account that the first character is flashed for 10 stimulus cycles, thus for each group of words the specific time T was computed following the parameters reported in Table 5.1.

The results listed in Table 5.2 show that even if the duration increased with the AC2* method with respect to the AC2 method with 3 and 4 repetitions, nevertheless the ITR increased thanks to the improvement of the average accuracy. Indeed the ITR is a parameter to evaluate BCI performance that takes in account both speed ($1/T$) and accuracy, in this case the benefit reached in term of accuracy by the AC2* method overcompensates the increase of time required to spell a character with respect to the AC2 method. On the other hand, we can notice that with 10 stimulus cycles the duration increased excessively and, although the average accuracy is higher, the ITR is lower with respect to the ITR computed for 3 and 4 repetitions and low for a real application.

TABLE 5.2 – Theoretical average ITR (bpm), computed for AC2 method for 3, 4 and 10 stimulus cycles (s.c.) and AC2* method for 3 and 4 stimulus cycles (s.c.).

	Session 1				
	3 s.c.		4 s.c.		10 s.c.
	AC2	AC2*	AC2	AC2*	AC2
3-char	25.1	27.6	25.4	27.1	17.3
4-char	16.7	23.3	21.8	26.4	20.7
5-char	15.8	28.9	15.1	28.2	20.1
sentence	12.7	22.1	16.4	20.3	15.8

	Session 2				
	3 s.c.		4 s.c.		10 s.c.
	AC2	AC2*	AC2	AC2*	AC2
3-char	26.6	33.1	26.6	35.7	18.5
4-char	21.1	25.9	21.1	28.5	18
5-char	20.3	29.6	20.3	29.7	16.7
sentence	6.7	13.3	6.7	11	16

The average theoretical ITR, listed in Table 5.2, demonstrates that both proposed methods, AC2 and AC2* with 3 or 4 stimulus cycles, could compete with the state-of-art of BCI speller with language model [Mora-Cortes et al., 2014; Speier et al., 2016]. Furthermore, considering for example a subject who reaches an accuracy of 100 % in the spelling of a sentence (Figure 5.5, Session 1), we can obtain with the AC2* method an ITR of around 54 bpm with 3 stimulus cycles and around 52 bpm with 4 stimulus cycles. These theoretical results can effectively compete with, for example, the work proposed by [Gembler and Volosyak, 2019], in which they reported an average ITR value of 57 bpm for the spelling of a sentence in a c-VEP BCI.

5.4.2 Further remarks

The proposed auto-calibration is one component contributing to the performance of c-VEP. At least two more improvements should be made. First, the language model should be integrated into an online application, including specific features such as error correction and word completion [Speier et al.,

2016]. Second, different strategies may improve the repeatability of the VEP responses, which is crucial for the performance of the method. One way is to investigate the influence of the stimulus modulation in order to design an appropriate interface for the auto-calibration c-VEP BCI system. Indeed the stimulus modulation is crucial to obtain a high performance c-VEP BCI. Many studies [Wei et al., 2016; Isaksen et al., 2017; Turi and Clerc, 2019] investigated several parameters to understand their influence in a c-VEP system, but analyzing these studies is clear that it is not possible to define an universal optimal stimulus parameter setting suitable for each BCI user in a c-VEP BCI system. For sure, the design of the interface in terms of optimal parameters of stimulation directly influences the usability of the system.

In addition, it will be useful to give to the user a feedback during the spelling. The GUI of the speller could provide step by step the possible partial or full words proposed by the system. In this way the user could correct the spelling of a wrong word by a specific character for this purpose (i.e. $'<'$) included in the virtual keyboard. Of course a modification of the dictionary would be required. Indeed, it will be necessary to add duplicated versions of each word belonging to the dictionary and each possible "root" of each word will be followed by the specific character corresponding to the error correction command. For instance for the word "$SUN\_$" the possible sequences of characters that will be added to the dictionary are: "S<", "SU<", "SUN<". Practically, all the possible partial or full words will be shown on the screen to the user during the spelling and if the root of the target word is not belonging to the propositions, the user has to spell the specific character $'<'$. Thus, the system can recognize the error, delete the spelling in progress and the user can restart the spelling from the beginning of the word.

Moreover, a "new word mode" could be integrated in a real application of the system in order to overcome the limiting factor introduced by the use of a dictionary in the speller. When the user needs to spell a word not present in the dictionary, the "new word mode" could be entered by spelling a specific sequence of characters defined during the design of the system (e.g the escape sequence ",?_"). In this mode the speller would allow to insert a new word character by character, with each character followed by the "_" symbol. For example, to type a password the user would enter the "new word mode" and spell "1_3_6_7_", finally going back to the normal mode using the same escape sequence. Then the system could also automatically add the

new word to the dictionary, so that the user can utilize the speller with the auto-calibration method without any restriction.

5.5 Conclusion

The method presented in our work is a proof of concept of auto-calibration for a c-VEP BCI speller. The method exploits only prior language information and not prior information from other training subjects. The intuition of the method is to use the fundamental property that characterizes the VEP response elicited from pseudo-random stimulus sequence. The response of each character is identical, but circularly-shifted with a specific time sample lag. The method is based on the extraction of the time lag of the VEP response of each character with respect to the VEP of the first character during the spelling of the target word. After the spelling of each character, the system finds within the dictionary all the possible words whose letters respect the sequence of relative lags. Offline experiments were simulated to test the proposed auto-calibration method compared to the standard calibration method. We demonstrated that our method can compete with the current state-of-art of BCI spellers [Speier et al., 2016; Gembler and Volosyak, 2019].

Furthermore, the proposed method shows promising results to develop a calibration-free c-VEP BCI. This opens also a new perspectives to the diffusion of BCI systems more user-friendly and adaptable to each user.

Part V

Conclusion

6 Conclusion and Perspectives

The purpose of BCIs is to restore communication skills for people with severe motor disabilities, but many limitations [Lotte et al., 2015] prevent the diffusion of BCI systems in real applications, outside the research laboratory. Among these limitations there is the tedious calibration phase and the lack of robustness of BCI systems. This thesis aims to propose novel strategies to develop an user-centered system in order to provide possible solutions to these limitations.

We started providing a review of the state-of-art of the work on the ERP and VEP spellers and, briefly, the SMR-BCIs, analyzing the main characteristics of these systems and also their limits. After an overview on these systems, we concluded that all require a user-centered development to face some of these limitations, such as the calibration stage, which is due to inter-subject and inter-session variability.

Many aspects of the system can be investigated to make these applications user-adaptive, in this manuscript the work was focused on the aspect of the calibration phase, proposing novel strategies to improve or avoid the traditional calibration phase.

In the following sections we summarize the contributions proposed in this thesis, their main limitations and possible improvements.

User-centered system for a c-VEP BCI speller

The first contribution proposed in this thesis is related to the development of user-centered systems for a c-VEP BCI speller and for a MI-BCI, presented and discussed in Chapter 4.

The goal of this contribution was to propose and evaluate different approaches to develop a system that can consider the central role of the user in different applications.

Adaptive parameter setting in c-VEP BCI

In Section 4.1 some BCI speller applications are presented describing their pipeline and referring to the state-of-the-art, in order to show the method mostly used in existing implementations. A particular focus was put to highlight the importance of the GUI in a speller, analyzing key parameters of each stimulation paradigm. Something that all such paradigms share is that there is no set of parameters universally valid for any of them, because there are many factors influencing the cerebral response, the most important one being the condition of the user. The lack of a set of universal stimulus parameters impacts the diffusion of the VEP spellers outside of the research laboratory, also because each speller has its own stimulus design, and consequently the research is concentrated on different optimizations for each type of speller.

A confirmation of this is discussed in Chapter 3, where we presented our pilot study for the development of a c-VEP BCI from scratch. We noticed that some stimulus parameters, such as the arrangement of the keyboard with borders or not, provided different results between our work and the others proposed in literature [Bin et al., 2011]. For this reason, we decided to investigate the parameters of stimulation to obtain a comfortable system for the user. The objective was to deploy a strategy to find the best stimulus parameters for each subject, in terms of comfort and performance. To reach this result, we replaced the long standard calibration phase with a shorter adaptive parameter setting phase.

We investigated four frequencies of stimulation (15 Hz, 20 Hz, 30 Hz, 60 Hz), proposing a different investigation approach with respect to the literature. Indeed, many works investigated higher stimulation frequencies [Nezamfar et al., 2015; Wittevrongel et al., 2017] in order to make the stimulation faster, to improve the performance of the system in terms of bit rate. In our case, we decided to analyze some lower stimulation frequencies to evaluate whether they were more comfortable for the subject and, consequently, improved the performance of our system, in terms of accuracy. We tested and evaluated the method on nine healthy subjects. The experiments consisted

in an adaptive parameter setting phase followed by a second phase in which the subject focused on imposed characters, in which each target flashed according to the set of parameters found during the adaptive phase. Then the acquired data were processed offline. The adaptive phase allowed to detect the best frequency of stimulation computing λ score, that can be considered as a performance estimator for a c-VEP BCI system, and allowed to estimate the influence of different frequencies of stimulation on the VEP responses.

The obtained results proved that our method does not require a long calibration phase such as other system in literature [Bin et al., 2011; Wei et al., 2016]. In our case we developed a method that can reach high values of accuracy defining the most appropriate stimulus for each subject, reaching good performance for many subjects. Moreover we demonstrated that while decreasing the flashing pattern frequency, it is possible to compensate the lengthening of the experiment by setting a lower number of stimulus cycles, showing that stimulus duration at different frame rates does not impact the performance of the system. Furthermore, the achieved results shows that for our system the subjects preferred low stimulation frequencies rather that the standard stimulation frequency of 60 Hz or higher, as reported in the work of [Borhanazad et al., 2019]. This confirms how much each system is independent of the others in the setting of stimulation parameters.

Nevertheless, this study lacks in a comparison with a non-adaptive system, consequently it is not possible to quantify the real improvement in terms of performance with respect to a standard c-VEP system. This lack was caused by the inexperience in the field of experimental protocol design and by the difficult to recruit volunteers in the experimentation.

Another limitation of the proposed method is that, for the moment, it was not tested online. We would like to underline that the proposed work could pave the way for the development of an user-centered system but many improvements can be made. For instance combining different parameters of stimulation in the adaptive setting phase and applying this phase to model the VEP response characteristic to each subject, independently of the session. Moreover, the spelling phase can be improved including an early stopping strategy [Thomas et al., 2014; Thielen et al., 2015], in such a way to limit the number of stimulation cycles, reducing the stimulus duration. Optimally, these results would lead to the design of a subject-dependent c-VEP BCI with

high communication performance.

Design of MI-BCI system for a motor-impaired user

In Section 4.2 we proposed a long multi-stage training for a motor-impaired user in the context of the Cybathlon BCI competition. Even for a SMR-BCI system it was necessary to focus the work on the calibration phase to build an user-centered system, but in a different way with respect to the c-VEP BCI. Indeed, in this case the training phase was fundamental to train the user to control his/her mental tasks in a performing way.

So we mainly focused our work on the design of a specific experimental protocol for our pilot, in order to define the mental tasks more suitable for the pilot and also with the aim to train her in the most efficient way possible, in order to be competitive for a BCI-race. We proposed a multi-stage training protocol designed specifically for our pilot. This protocol took two phases, the investigation phase and training phase, for a total duration of 3 months. The results reached during the training phase showed the ability of the subject to control some tasks better than others across sessions. Moreover, analyzing the results, it is evident that the control of a 4-class paradigm required more concentration and ability than a 3-class paradigm. This aspect is also reported in many works in literature [Ponferrada et al., 2018], in which it is often highlighted the need of systems adaptive to the user, in particular for MI-BCI [Saha and Baumert, 2019].

Thanks the participation to the Cybathlon BCI series we tested our system in a real competition, comprising the limits of these types of system and the factors that influence the performance, such as the high concentration required for the pilot to be competitive. Moreover, we noticed a great degree of variability between pilots. For instance, motor abilities were highly variable from one pilot to another and we understand that the role of the user is fundamental in a SMR-BCI system too, confirming the need to develop user-centered system, in particular for disabled people who present different needs. This highlights the importance of developing a personalized BCI system, to tackle the needs of each user in any situation, such as a competition or real life.

Auto-calibration of c-VEP BCI by word prediction

The second contribution deals with the development of a novel auto-calibration-based word prediction c-VEP BCI speller, presented in Chapter 5.

The main objective of this method was to avoid the long and annoying calibration phase, allowing to spell words exploiting the main features that characterized the VEP response and the language information. The proposed model extracts relative lags between successive characters from the VEP responses and then predicts the full word using a dictionary, eliminating the traditional calibration phase. To the best of our knowledge, this is the first work proposing a word-prediction-based auto-calibration method for c-VEP BCIs.

We evaluated the method on a public dataset. We simulated the spelling of four different groups of 5 words, in order to evaluate the performance of the method with a different number of characters. We conducted experiments applying the standard calibration (C), the auto-calibration method considering only the most correlated character (AC1) and the auto-calibration method considering the two most correlated characters (AC2). Each experiment was parameterized by 3, 4, 10, 15 and 20 stimulus cycles across sessions. The results reached comparing the different methods showed that the average accuracy improved when increasing the number of stimulus cycles, for all methods. Furthermore, the AC2 method performed better than the AC1 method for all the tested cases.

So, we simulated others experiments proposing an option of the AC2 method, called AC2* method, in which we evaluated the performance of AC2 method to spell the same groups of words with a number of stimulus cycles fixed to 10 for the first character and following the parameterized number of stimulus cycles for all the other characters belonging to each word.

The achieved results showed that the average accuracy increases with the AC2* method, reaching values always higher than the ones obtained with the AC2 method, usually close to the accuracy reached with 10 stimulus cycles, denoting a particular benefit for the performance with a lower number of stimulus cycles. In order to compare the effectiveness of the proposed method with the state-of-art, the theoretical ITR was computed. The results demonstrated that both proposed methods, AC2 and AC2*, could compete

with the state-of-art of BCI speller with language model [Mora-Cortes et al., 2014; Speier et al., 2016; Gembler and Volosyak, 2019].

The proposed method is a proof of concept, that can contribute to the performance of c-VEP BCI spellers, however more improvements should be made for an online application. The language model should be integrated into the system, including specific features such as error correction and word completion. Moreover, different strategies may improve the repeatability of the VEP responses, which is crucial for the performance of the proposed method. One way is to investigate the influence of the stimulus modulation in order to design an appropriate interface for the auto-calibration c-VEP BCI system. Indeed the stimulus modulation is crucial to obtain a high performance c-VEP BCI and, as discussed in Section 2.1 and in the previous contribution (Section 4.1), nowadays it is not possible to define an universal optimal stimulus parameter setting suitable for each BCI user in a c-VEP BCI system.

Further remarks and future perspectives

Several open questions are raised by the work presented in this thesis about the needs and strategies to develop user-centered system.

Most of this manuscript has been focused on the study of BCI spellers, especially c-VEP spellers. Choosing to develop a system from scratch and testing it allowed us to identify some aspects that would need to be investigated and considered more in the BCI research. Among them there is the GUI standardization, whose advancement could effectively improve this type of system, which is very promising but still limited in diffusion in the BCI research community and applications.

Another aspect that should be considered is that the effectiveness of the system is influenced by the quality of the data, specifically of the VEP response. Many works proposed spatial filters and classification algorithms to improve the detection of the VEP response. However nowadays a few works propose [Nagel and Spüler, 2018; Thielen et al., 2015] a generative model valid for the VEP response. Certainly, the definition of a model for the c-VEP response could improve the diffusion and research based on this system, that has demonstrated to reach very high performance [Bin et al., 2011]. Indeed, a valid model of the VEP response would improve detection during

the spelling phase but above all it would allow to perform a greater number of simulations offline, since the datasets currently available are really scarce, opening new research directions on c-VEP BCI spellers.

To conclude, we think that this manuscript highlights the needs to focus the research on the development of systems more centered on the user, in order to develop systems applicable in daily life. With this goal, we proposed novel strategies that can be a starting point for the development of such user-friendly BCI systems.

List of Publications

Journal Papers

— F. Turi, N. Gayraud, M. Clerc. "Auto-calibration for c-VEP BCI by word prediction" (*Paper submitted*)

— F. Turi, A. Audino, M. Clerc, T. Papadopoulo. "Long multi-stage training for a motor-impaired user in a BCI competition" (*Paper in preparation*)

Participation in Conferences

— F. Turi, M. Clerc. "Adaptive Stimulus Parameter Setting for cVEP BCI", *Proceedings of the 8th Graz Brain-Computer Interface Conference 2019*. Graz, Austria. (*Paper accepted*)

— F. Turi, M. Clerc. "Design of a subject-dependent cVEP brain-computer interface". *First meeting NeuroMod Institute 2019*, Frejus, France. (*Oral presentation*)

— F. Turi, M. Clerc. "Adaptive Stimulus Parameter Setting for cVEP BCI", *Journée Jeunes Chercheurs en Interface Cerveau Ordinateur et Neurofeedback*, *JJC-ICON-2019*, Lille, France.

— F. Turi, N. Gayraud, M. Clerc. "Zero-calibration cVEP BCI using word prediction: a proof of concept", *Journée Jeunes Chercheurs en Interface Cerveau Ordinateur et Neurofeedback*, *JJC-ICON-2018*, Toulose, France. (*Oral presentation*)

— F. Turi, N. Gayraud, M. Clerc. "Zero-calibration cVEP BCI using word prediction: a proof of concept", *BCI 2018 - 7th International BCI Meeting*, Pacific Grove, California, United States.

— F. Turi, M. Clerc. "Stimuli strategies for VEP BCI", *Journée Jeunes Chercheurs en Interface Cerveau Ordinateur et Neurofeedback*, *JJC-ICON-2017*, Bordeaux, France.

— F. Turi, M. Clerc. "Comparison between P300 speller and cVEP BCI", *Rencontre C@UCA 2017*, Frejus, France.

Other experiences

— MOMI 2019 – Le Monde des Mathématiques Industrielles, February 2019, INRIA Sophia Antipolis, France.
— CoBCoM 2017 – Computational Brain Connectivity Mapping, Winter School Workshop, 20–24 November 2017, Juan Les Pins, France.
— MOMI 2017 – Le Monde des Mathématiques Industrielles, February 2017, INRIA Sophia Antipolis, France.
— Summer School on Brain Connectomics, 19–22 September 2016, University of Verona, Italy.
— Temporary professor in Signal and Image Processing at Polytech of Sophia, 2017–2020.

Award

— Student Award BCI Meeting 2018, California, United States.

References

Bibliography

Abdulkader, S. N., Atia, A., and Mostafa, M.-S. M. (2015). "Brain computer interfacing: Applications and challenges". In: *Egyptian Informatics Journal* 16.2, pp. 213–230.

Aggarwal, S. and Chugh, N. (2019). "Signal processing techniques for motor imagery brain computer interface: A review". In: *Array* 1, p. 100003.

Allison, B. Z. and Pineda, J. A. (2003). "ERPs evoked by different matrix sizes: implications for a brain computer interface (BCI) system". In: *IEEE transactions on neural systems and rehabilitation engineering* 11.2, pp. 110–113.

Aminaka, D., Makino, S., and Rutkowski, T. M. (2015). "Chromatic and high-frequency cVEP-based BCI paradigm". In: *2015 37th EMBC*. IEEE, pp. 1906–1909.

Barbosa, A. O., Achanccaray, D. R., and Meggiolaro, M. A. (2010). "Activation of a mobile robot through a brain computer interface". In: *2010 IEEE International Conference on Robotics and Automation*. IEEE, pp. 4815–4821.

Benabid, A. L. et al. (2019). "An exoskeleton controlled by an epidural wireless brain–machine interface in a tetraplegic patient: a proof-of-concept demonstration". In: *The Lancet Neurology* 18.12, pp. 1112–1122.

Bieger, J., Molina, G. G., and Zhu, D. (2010). "Effects of stimulation properties in steady-state visual evoked potential based brain-computer interfaces". In: *Proceedings of 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Engineering in Medicine and Biology Society, pp. 3345–8.

Bin, G. et al. (2009a). "VEP-based brain-computer interfaces: time, frequency, and code modulations". In: *IEEE Computational Intelligence Magazine* 4.4.

Bin, G. et al. (2011). "A high-speed BCI based on code modulation VEP". In: *Journal of neural engineering* 8.2, p. 025015.

Bin, G. et al. (2009b). "An online multi-channel SSVEP-based brain–computer interface using a canonical correlation analysis method". In: *Journal of neural engineering* 6.4, p. 046002.

Birbaumer, N. and Cohen, L. G. (2007). "Brain–computer interfaces: communication and restoration of movement in paralysis". In: *The Journal of physiology* 579.3, pp. 621–636.

Blankertz, B. et al. (2010). "The Berlin brain–computer interface: non-medical uses of BCI technology". In: *Frontiers in neuroscience* 4, p. 198.

Borhanazad, M., Thielen, J., and Desain, P. (2019). "The effect of high and low frequencies in C-VEP BCI". In: *8th Graz Brain-Computer Interface Conference*.

Britton, J. W. et al. (2016). *Electroencephalography (EEG): An introductory text and atlas of normal and abnormal findings in adults, children, and infants*. American Epilepsy Society, Chicago.

Cao, T. et al. (2012). "Flashing color on the performance of SSVEP-based brain-computer interfaces". In: *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, pp. 1819–1822.

Cecotti, H. (2010). "A self-paced and calibration-less SSVEP-based brain–computer interface speller". In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 18.2, pp. 127–133.

Cecotti, H. and Rivet, B. (2010). "One step beyond rows and columns flashes in the p300 speller: a theoretical description". In: *International Journal of bioelectromagnetism* 13.1, pp. 39–41.

Cheng, M. et al. (2002). "Design and implementation of a brain-computer interface with high transfer rates". In: *IEEE transactions on biomedical engineering* 49.10, pp. 1181–1186.

Cincotti, F. et al. (2008). "Non-invasive brain–computer interface system: towards its application as assistive technology". In: *Brain research bulletin* 75.6, pp. 796–803.

Coyle, S. et al. (2004). "On the suitability of near-infrared (NIR) systems for next-generation brain–computer interfaces". In: *Physiological measurement* 25.4, p. 815.

Daly, J. J. and Wolpaw, J. R. (2008). "Brain–computer interfaces in neurological rehabilitation". In: *The Lancet Neurology* 7.11, pp. 1032–1043.

Donchin, E. and Coles, M. G. (1988). "Is the P300 component a manifestation of context updating?" In: *Behavioral and brain sciences* 11.3, pp. 357–374.

Evain, A. (2016). "Optimizing the use of SSVEP-based brain-computer interfaces for human-computer interaction". PhD dissertation. Université Rennes 1.

Farwell, L. A. and Donchin, E. (1988). "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials". In: *Electroencephalography and clinical Neurophysiology* 70.6, pp. 510–523.

Gao, Q. et al. (2019). "Channel Projection-Based CCA Target Identification Method for an SSVEP-Based BCI System of Quadrotor Helicopter Control". In: *Computational Intelligence and Neuroscience* 2019.

Gayraud, N. (2018). "Adaptive machine learning methods for event related potential-based brain computer interfaces". PhD thesis.

Gembler, F. and Volosyak, I. (2019). "A Novel Dictionary-Driven Mental Spelling Application Based on Code-Modulated Visual Evoked Potentials". In: *Computers* 8.2, p. 33.

Gembler, F. W. et al. (2020). "Five shades of grey: exploring quintary m-sequences for more user-friendly c-VEP-based BCIs". In: *Computational Intelligence and Neuroscience* 2020.

Gold, R. (1967). "Optimal binary sequences for spread spectrum multiplexing". In: *IEEE Transactions on information theory* 13.4, pp. 619–621.

Golomb, S. W. (1982). *Shift Register Sequences*. Aegean Park Press, Laguna Hill.

Gonsalvez, C. J. and Polich, J. (2002). "P300 amplitude is determined by target-to-target interval". In: *Psychophysiology* 39.3, pp. 388–396.

Guger, C. et al. (2012). "Comparison of dry and gel based electrodes for P300 brain–computer interfaces". In: *Frontiers in neuroscience* 6, p. 60.

Guy, V. et al. (2018). "Brain computer interface with the P300 speller: usability for disabled people with amyotrophic lateral sclerosis". In: *Annals of physical and rehabilitation medicine* 61.1, pp. 5–11.

Hamadicharef, B. (2010). "AUC confidence bounds for performance evaluation of Brain-Computer Interface". In: *2010 3rd International Conference on Biomedical Engineering and Informatics*. Vol. 5. IEEE, pp. 1988–1991.

He, B. et al. (2013). "Brain–computer interfaces". In: *Neural engineering*. Springer, pp. 87–151.

Hoffmann, U. et al. (2008). "An efficient P300-based brain–computer interface for disabled subjects". In: *Journal of Neuroscience methods* 167.1, pp. 115–125.

Hübner, D. et al. (2017). "Learning from label proportions in brain-computer interfaces: Online unsupervised learning with guarantees". In: *PloS one* 12.4, e0175856.

Huebner, D. et al. (2018). "Unsupervised learning for brain-computer interfaces based on event-related potentials: Review and online comparison". In: *IEEE Computational Intelligence Magazine* 13.2, pp. 66–77.

Isaksen, J. L., Mohebbi, A., and Puthusserypady, S. (2017). "Optimal pseudorandom sequence selection for online c-VEP based BCI control applications". In: *PloS one* 12.9, e0184785.

Islam, M. K., Rastegarnia, A., and Yang, Z. (2016). "Methods for artifact detection and removal from scalp EEG: A review". In: *Neurophysiologie Clinique/Clinical Neurophysiology* 46.4-5, pp. 287–305.

Jayaram, V. et al. (2016). "Transfer learning in brain-computer interfaces". In: *IEEE Computational Intelligence Magazine* 11.1, pp. 20–31.

Jia, C. et al. (2010). "Frequency and phase mixed coding in SSVEP-based brain–computer interface". In: *IEEE Transactions on Biomedical Engineering* 58.1, pp. 200–206.

Jiang, X., Bian, G., and Tian, Z. (2019). "Removal of artifacts from EEG signals: a review". In: *Sensors* 19.5, p. 987.

Jin, J. et al. (2012). "The changing face of P300 BCIs: a comparison of stimulus changes in a P300 BCI involving faces, emotion, and movement". In: *PloS one* 7.11.

Jin, J. et al. (2013). "Whether generic model works for rapid ERP-based BCI calibration". In: *Journal of neuroscience methods* 212.1, pp. 94–99.

Jukiewicz, M. and Cysewska-Sobusiak, A. (2016). "Stimuli design for SSVEP-based brain computer-interface". In: *International Journal of Electronics and Telecommunications* 62.2, pp. 109–113.

Kanwisher, N. G. (1987). "Repetition blindness: Type recognition without token individuation". In: *Cognition* 27.2, pp. 117–143.

Kaufmann, T. et al. (2012). "Spelling is just a click away–a user-centered brain-computer interface including auto-calibration and predictive text entry". In: *Frontiers in neuroscience* 6, p. 72.

Kauhanen, L. et al. (2007). "EEG-based brain-computer interface for tetraplegics". In: *Computational intelligence and neuroscience* 2007.

Kindermans, P.-J., Verstraeten, D., and Schrauwen, B. (2012a). "A bayesian model for exploiting application constraints to enable unsupervised training of a P300-based BCI". In: *PloS one* 7.4, e33758.

Kindermans, P. J. et al. (2012b). "A P300 BCI for the masses: Prior information enables instant unsupervised spelling". In: *Advances in Neural Information Processing Systems*, pp. 710–718.

Kleih, S. C. and Kübler, A. (2015). "Psychological factors influencing brain-computer interface (BCI) performance". In: *SMC, 2015 IEEE Intern. Conf.* Pp. 3192–3196.

Klem, G. H. (1999). "The ten-twenty electrode system of the international federation. The internanional federation of clinical nenrophysiology." In: *Electroencephalogr. Clin. Neurophysiol. Suppl.* 52, pp. 3–6.

Kübler, A. et al. (2001). "Brain-computer communication: self-regulation of slow cortical potentials for verbal communication". In: *Archives of physical medicine and rehabilitation* 82.11, pp. 1533–1539.

Kumar, J. S. and Bhuvaneswari, P. (2012). "Analysis of Electroencephalography (EEG) signals and its categorization–a study". In: *Procedia engineering* 38, pp. 2525–2536.

Lee, P. L. et al. (2008). "Brain computer interface using flash onset and offset visual evoked potentials". In: *Clinical Neurophysiology* 119.3, pp. 605–616.

Leeb, R. et al. (2007). "Self-paced (asynchronous) BCI control of a wheelchair in virtual environments: a case study with a tetraplegic". In: *Computational intelligence and neuroscience* 2007.

Libenson, M. H. (2012). *Practical Approach to Electroencephalography*. Elsevier Health Sciences.

Lin, Z. et al. (2006). "Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs". In: *IEEE transactions on biomedical engineering* 53.12, pp. 2610–2614.

Linden, D. E. (2005). "The P300: where in the brain is it produced and what does it tell us?" In: *The Neuroscientist* 11.6, pp. 563–576.

Lotte, F. (2014). "A tutorial on EEG signal-processing techniques for mental-state recognition in brain-computer interfaces". In: *Guide to Brain-Computer Music Interfacing*. Springer, pp. 133–161.

Lotte, F. (2015). "Signal processing approaches to minimize or suppress calibration time in oscillatory activity-based brain-computer interfaces". In: *Proceedings of the IEEE* 103.6, pp. 871–890.

Lotte, F. et al. (2007). "A review of classification algorithms for EEG-based brain–computer interfaces". In: *Journal of neural engineering* 4.2, R1.

Lotte, F., Bougrain, L., and Clerc, M. (2015). "Electroencephalography (EEG)-Based Brain–Computer Interfaces". In: *Wiley Encyclopedia of Electrical and Electronics Engineering*, pp. 1–20.

Lotte, F. et al. (2018). "A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update". In: *Journal of neural engineering* 15.3, p. 031005.

Lu, J. et al. (2013). "The effects of stimulus timing features on P300 speller performance". In: *Clinical Neurophysiology* 124.2, pp. 306–314.

Lu, S., Guan, C., and Zhang, H. (2009). "Unsupervised brain computer interface based on intersubject information and online adaptation". In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 17.2, pp. 135–145.

Materka, A. and Poryzała, P. (2014). "A robust asynchronous ssvep brain-computer interface based on cluster analysis of canonical correlation coefficients". In: *Human-Computer Systems Interaction: Backgrounds and Applications 3*. Springer, pp. 3–14.

Mellinger, J. et al. (2007). "An MEG-based brain–computer interface (BCI)". In: *Neuroimage* 36.3, pp. 581–593.

Middendorf, M. et al. (2000). "Brain-computer interfaces based on the steady-state visual-evoked response". In: *IEEE transactions on rehabilitation engineering* 8.2, pp. 211–214.

Millán, J. d. R. et al. (2010). "Combining brain–computer interfaces and assistive technologies: state-of-the-art and challenges". In: *Frontiers in neuroscience* 4, p. 161.

Mora-Cortes, A. et al. (2014). "Language model applications to spelling with brain-computer interfaces". In: *Sensors* 14.4, pp. 5967–5993.

Müller-Putz, G. R. et al. (2005). "Steady-state visual evoked potential (SSVEP)-based communication: impact of harmonic frequency components". In: *Journal of neural engineering* 2.4, p. 123.

Nagel, S. and Spüler, M. (2018). "Modelling the brain response to arbitrary visual stimulation patterns for a flexible high-speed Brain-Computer Interface". In: *PloS one* 13.10, e0206107.

Nakanishi, M. et al. (2017). "Enhancing detection of SSVEPs for a high-speed brain speller using task-related component analysis". In: *IEEE Transactions on Biomedical Engineering* 65.1, pp. 104–112.

Naseer, N. and Hong, K. S. (2015). "fNIRS-based brain-computer interfaces: a review". In: *Frontiers in human neuroscience* 9, p. 3.

Nezamfar, H., Salehi, S., and Erdogmus, D (2015). "Stimuli with opponent colors and higher bit rate enable higher accuracy for C-VEP BCI". In: *2015 SPMB*. IEEE, pp. 1–6.

Nicolas-Alonso, L. F. and Gomez-Gil, J. (2012). "Brain computer interfaces, a review". In: *Sensors* 12.2, pp. 1211–1279.

Oikonomou, V. P. et al. (2016). "Comparative evaluation of state-of-the-art algorithms for SSVEP-based BCIs". In: *arXiv preprint arXiv:1602.00904*.

Orhan, U. et al. (2012). "RSVP keyboard: An EEG based typing interface". In: *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, pp. 645–648.

Ostry, D. J. and Gribble, P. L. (2016). "Sensory plasticity in human motor learning". In: *Trends in neurosciences* 39.2, pp. 114–123.

Perrin, M. et al. (2012). "Objective and subjective evaluation of online error correction during P300-based spelling". In: *Advances in Human-Computer Interaction* 2012.

Pfurtscheller, G. and Da Silva, F. L. (1999). "Event-related EEG/MEG synchronization and desynchronization: basic principles". In: *Clinical neurophysiology* 110.11, pp. 1842–1857.

Ponferrada, E. G., Sylaidi, A., and Faisal, A. A. (2018). "Data-efficient Motor Imagery Decoding in Real-time for the Cybathlon Brain-Computer Interface Race". In:

Quadrianto, N. et al. (2009). "Estimating labels from label proportions." In: *Journal of Machine Learning Research* 10.10.

Raymond, J. E., Shapiro, K. L., and Arnell, K. M. (1992). "Temporary suppression of visual processing in an RSVP task: An attentional blink?" In: *Journal of experimental psychology: Human perception and performance* 18.3, p. 849.

Renard, Y. et al. (2010). "Openvibe: An open-source software platform to design, test, and use brain–computer interfaces in real and virtual environments". In: *Presence: teleoperators and virtual environments* 19.1, pp. 35–53.

Rezeika, A. et al. (2018). "Brain–computer interface spellers: A review". In: *Brain sciences* 8.4, p. 57.

Riener, R. (2016). "The Cybathlon promotes the development of assistive technology for people with physical disabilities". In: *Journal of neuroengineering and rehabilitation* 13.1, p. 49.

Rivet, B. et al. (2009). "xDAWN algorithm to enhance evoked potentials: application to brain–computer interface". In: *IEEE Transactions on Biomedical Engineering* 56.8, pp. 2035–2043.

Ryan, D. B. et al. (2010). "Predictive spelling with a P300-based brain–computer interface: increasing the rate of communication". In: *Intl. Journal of Human–Computer Interaction* 27.1, pp. 69–84.

Saha, S. and Baumert, M. (2019). "Intra-and inter-subject variability in EEG-based sensorimotor brain computer interface: a review". In: *Frontiers in Computational Neuroscience* 13, p. 87.

Samizo, E., Yoshikawa, T., and Furuhashi, T. (2013). "A study on application of RB-ARQ considering probability of occurrence and transition probability for P300 speller". In: *International Conference on Augmented Cognition*. Springer, pp. 727–733.

Schalk, G. and Leuthardt, E. C. (2011). "Brain-computer interfaces using electrocorticographic signals". In: *IEEE reviews in biomedical engineering* 4, pp. 140–154.

Sellers, E. W. et al. (2006). "A P300 event-related potential brain–computer interface (BCI): the effects of matrix size and inter stimulus interval on performance". In: *Biological psychology* 73.3, pp. 242–252.

Serby, H., Yom-Tov, E., and Inbar, G. F. (2005). "An improved P300-based brain-computer interface". In: *IEEE Transactions on neural systems and rehabilitation engineering* 13.1, pp. 89–98.

Singla, R. and Haseena, B. (2014). "Comparison of ssvep signal classification techniques using svm and ann models for bci applications". In: *International Journal of Information and Electronics Engineering* 4.1, p. 6.

Sokolova, M. and Lapalme, G. (2009). "A systematic analysis of performance measures for classification tasks". In: *Information processing and management* 45.4, pp. 427–437.

Speier, W. et al. (2011). "Natural language processing with dynamic classification improves P300 speller accuracy and bit rate". In: *Journal of neural engineering* 9.1, p. 016004.

Speier, W., Arnold, C., and Pouratian, N. (2013). "Evaluating true BCI communication rate through mutual information and language models". In: *PLoS One* 8.10.

Speier, W., Arnold, C., and Pouratian, N. (2016). "Integrating language models into classifiers for BCI communication: a review". In: *Journal of neural engineering* 13.3, p. 031002.

Spüler, M., Rosenstiel, W., and Bogdan, M. (2012). "Online adaptation of a c-VEP brain-computer interface (BCI) based on error-related potentials and unsupervised learning". In: *PloS one* 7.12, e51077.

Spüler, M. et al. (2014). "Spatial filtering based on canonical correlation analysis for classification of evoked or event-related potentials in EEG data". In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 22.6, pp. 1097–1103.

Sur, S. and Sinha, V. (2009). "Event-related potential: An overview". In: *Industrial psychiatry journal* 18.1, p. 70.

Sutter, E. E. (1992). "The brain response interface: communication through visually-induced electrical brain responses". In: *Journal of Microcomputer Applications* 15.1, pp. 31–45.

Swets, J. A. (1988). "Measuring the accuracy of diagnostic systems". In: *Science* 240.4857, pp. 1285–1293.

Tanaka, K., Matsunaga, K., and Wang, H. O. (2005). "Electroencephalogram-based control of an electric wheelchair". In: *IEEE transactions on robotics* 21.4, pp. 762–766.

Tatum IV, W. O. (2018). *Atlas of Artifacts in Clinical Neurophysiology*. Springer Publishing Company.

Tello, R. J.M. G. et al. (2015). "Comparison of the influence of stimuli color on steady-state visual evoked potentials". In: *Research on Biomedical Engineering* AHEAD, pp. 0–0.

Thielen, J. et al. (2015). "Broad-Band visually evoked potentials: re (con) volution in brain-computer interfacing". In: *PloS one* 10.7, e0133797.

Thomas, E. et al. (2013). "Optimizing P300-speller sequences by RIP-ping groups apart". In: *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*. IEEE, pp. 1062–1065.

Thomas, E. et al. (2014). "CoAdapt P300 speller: optimized flashing sequences and online learning". In: *Sixth International Brain-Computer Interface Conference 2014*.

Townsend, G. et al. (2010). "A novel P300-based brain–computer interface stimulus presentation paradigm: moving beyond rows and columns". In: *Clinical neurophysiology* 121.7, pp. 1109–1120.

Turi, F. (2015). *A Brain-Computer Interface Based on Single-Trial Movement-Related Cortical Potentials*. Politecnico di Torino.

Turi, F. and Clerc, M. (2019). "Adaptive parameter setting in a code modulated visual evoked potentials BCI". In: *8th Graz Brain-Computer Interface Conference*.

Van Dokkum, L., Ward, T., and Laffont, I. (2015). "Brain computer interfaces for neurorehabilitation–its current status as a rehabilitation strategy post-stroke". In: *Annals of physical and rehabilitation medicine* 58.1, pp. 3–8.

Van Gerven, M. et al. (2009). "The brain–computer interface cycle". In: *Journal of neural engineering* 6.4, p. 041001.

Venables, L. and Fairclough, S. H. (2009). "The influence of performance feedback on goal-setting and mental effort regulation". In: *Motivation and Emotion* 33.1, pp. 63–74.

Verhoeven, T. et al. (2017). "Improving zero-training brain-computer interfaces by mixing model estimators". In: *Journal of neural engineering* 14.3, p. 036021.

Vialatte, F. B. et al. (2010). "Steady-state visually evoked potentials: focus on essential paradigms and future perspectives". In: *Progress in neurobiology* 90.4, pp. 418–438.

Volosyak, I. (2011). "SSVEP-based Bremen–BCI interface—boosting information transfer rates". In: *Journal of neural engineering* 8.3, p. 036020.

Volosyak, I., Cecotti, H., and Gräser, A. (2009). "Impact of frequency selection on LCD screens for SSVEP based brain-computer interfaces". In: *International Work-Conference on Artificial Neural Networks*. Springer, pp. 706–713.

Wei, Q., Feng, S., and Lu, Z (2016). "Stimulus specificity of brain-computer interfaces based on code modulation visual evoked potentials". In: *PloS one* 11.5, e0156416.

Wei, Q., Xiao, M., and Lu, Z. (2011). "A comparative study of canonical correlation analysis and power spectral density analysis for SSVEP detection". In: *2011 Third International Conference on Intelligent Human-Machine Systems and Cybernetics*. Vol. 2. IEEE, pp. 7–10.

Weiskopf, N. et al. (2003). "Physiological self-regulation of regional brain activity using real-time functional magnetic resonance imaging (fMRI): methodology and exemplary data". In: *Neuroimage* 19.3, pp. 577–586.

Wittevrongel, B., Van Wolputte, E., and Van Hulle, M. M. (2017). "Code-modulated visual evoked potentials using fast stimulus presentation and spatiotemporal beamformer decoding". In: *Scientific reports* 7.1, p. 15037.

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

Wolpaw, J. R. et al. (2000). "Brain-computer interface technology: a review of the first international meeting". In: *IEEE transactions on rehabilitation engineering* 8.2, pp. 164–173.

Wolpaw, J. R. et al. (2002). "Brain-computer interfaces for communication and control". In: *Clinical neurophysiology* 113.6, pp. 767–791.

Yang, J. et al. (2012). "Channel selection and classification of electroencephalogram signals: an artificial neural network and genetic algorithm-based approach". In: *Artificial intelligence in medicine* 55.2, pp. 117–126.

Yuan, H. and He, B. (2014). "Brain–computer interfaces using sensorimotor rhythms: current state and future perspectives". In: *IEEE Transactions on Biomedical Engineering* 61.5, pp. 1425–1435.

Zander, T. O. and Kothe, C. (2011). "Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general". In: *Journal of neural engineering* 8.2, p. 025005.

Zhang, Y. et al. (2014). "Frequency recognition in SSVEP-based BCI using multiset canonical correlation analysis". In: *International journal of neural systems* 24.04, p. 1450013.

Zhu, D. et al. (2010). "A survey of stimulation methods used in SSVEP-based BCIs". In: *Computational intelligence and neuroscience* 2010.