Design and Study of Interactive Systems based on Brain- Computer Interfaces and Augmented Reality

Hakim Si Mohammed

UNIVERSITE BRETAGNE LOIRE / MATHSTIC

INSA RENNES

THÈSE DE DOCTORAT DE

L'INSA RENNES
COMUE UNIVERSITÉ BRETAGNE LOIRE

ÉCOLE DOCTORALE N° 601
Mathématiques et Sciences et Technologies de l'Information et de la Communication
Spécialité : *Informatique*

Par

« Hakim SI MOHAMMED »

« Design and Study of Interactive Systems based on Brain-Computer Interfaces and Augmented Reality »

Thèse présentée et soutenue à Rennes, le 03 Décembre 2019
Equipe Hybrid, Inria/IRISA
Thèse N° : 19ISAR 27 / D19 - 27

Rapporteurs avant soutenance :

Mark BILLINGHURST	Professeur, University of South Australia
François CABESTAING	Professeur, Université de Lille

Composition du Jury :

Président :	Bruno ARNALDI	Professeur, INSA de Rennes
Examinateurs :	Mark BILLINGHURST	Professeur, University of South Australia
	François CABESTAING	Professeur, Université de Lille
	Reinhold SCHERER	Professeur, University of Essex
	Tanja SCHULTZ	Professeure, University Bremen
Dir. de thèse :	Anatole LÉCUYER	Directeur de Recherche, Inria Rennes
Encadrants de thèse :	Ferran ARGELAGUET	Chargé de Recherche, Inria Rennes
	Géry CASIEZ	Professeur, Université de Lille

Invité(s) :

Fabien LOTTE	Directeur de Recherche, Inria Bordeaux

"*N'admettez rien à priori si vous pouvez le vérifier.*"

Rudyard Kipling

"*Et tout élan est aveugle, s'il n'est pas guidé par le savoir.*
Et tout savoir est vain, s'il n'est pas accompagné de labeur.
Et tout labeur est futile, s'il n'est pas accompli avec amour.."

Gibran Khalil Gibran

A ceux qui ne sont plus là...

REMERCIEMENTS

Ce parcours de trois années de thèse qui s'achève est autant empli de moments forts qu'il est passé très vite. C'est l'occasion de remercier tous ceux qui ont, d'une manière ou d'une autre, contribué à ce résultat.

J'ai plaisir à remercier en premier lieu l'ensemble des membres de mon Jury : M. le Professeur Bruno Arnaldi, pour m'avoir fait l'honneur de le présider, mes rapporteurs, MM. les Professeurs François Cabestaing et Mark Billinghurst (qui a en particulier veillé très tard en visioconference depuis l'Australie durant ma soutenance), ainsi que mes examinateurs, Mme la Professeure Tanja Schultz, M. le Professeur Reinhold Scherer et M. le Directeur de Recherche Fabien Lotte. Tous ont accordé beaucoup de leur précieux temps à la lecture de mes travaux, et je leur sais particulièrement gré pour leurs commentaires erichissants, pertinents et si constructifs.

J'exprime mon immense gratitude à ceux sans qui rien de tout cela n'aurait été possible. Ceux qui m'ont avant tout offert leur confiance en acceptant de m'accueillir, de diriger mes travaux, et auprès de qui j'ai tant appris. S'il fallait d'un mot remercier mon Directeur de thèse, Anatole Lécuyer, je dirais simplement : « Merci pour tout, Anatole. Pour absolument tout. ». Merci de m'avoir fait confiance, merci d'avoir su tirer le meilleur de moi-même, merci pour tes nombreux conseils. Merci d'avoir été là dans tous les moments difficiles, de m'avoir toujours soutenu et d'avoir toujours eu les mots justes en toutes circonstances. Et pour tout dire, merci de ton amitié. J'ai énormément appris à tes côtés, tant sur le plan scientifique que sur le plan personnel, et je suis honoré et particulièrement fier d'avoir travaillé à tes côtés et sous ta direction.

Un très grand merci à mes encadrants, Ferran Argelaguet et à Géry Casiez, qui m'ont, durant ce parcours de thèse, fait bénéficier d'un encadrement fait à la fois de compréhension et d'exigence, mais toujours d'un soutien qui m'a porté sans cesse. Merci à toi, Ferran, d'avoir toujours été là pendant ces trois ans. Merci d'avoir cru en moi, même quand les circonstances ou les délais n'étaient pas favorables, et d'avoir en même temps toujours su être la voix de la raison, celle sans qui je me serais plus d'une fois enflammé. Merci pour ton aide au quotidien tant dans la rédaction que dans le développement et les idées. Merci pour tous ces soirs où tu as veillé pour relire, corriger et améliorer mes papiers (notamment une certaine semaine inoubliable au Japon). Je garderai toujours le souvenir de tes « petits warning » qui s'avéraient tellement indiqués à chaque fois. Merci à toi également, Géry, pour tous nos échanges plus enrichissants

les uns que les autres, pour tes idées toujours si constructives ainsi que pour ton enthousiasme. C'était toujours avec un immense plaisir que venais à Lille, et je garderai un souvenir marquant de nos échanges en ces occasions.

J'ai aussi beaucoup de plaisir à remercier M. Nicolas Roussel. Bien que trop court à mon goût, du fait de ton départ pour de nouvelles responsabilités au sein d'Inria, ton encadrement au début de mon parcours a eu un très fort impact sur ma façon d'aborder mon sujet de thèse. Ton approche pragmatique des BCI et ton expérience des IHM m'ont beaucoup aidé tout au long de mon travail, et je tiens beaucoup à t'exprimer ma gratitude.

Je tiens beaucoup à remercier à présent ceux sans qui je n'en serais pas là aujourd'hui. Ceux qui, en amont et aujourd'hui encore, ont fait et font de moi celui que je suis aujourd'hui : Ma famille. Merci à vous, Papa, Maman et Sam d'avoir été là à chaque fois que j'en ai eu besoin, et surtout quand je ne savais pas moi-même que j'en avais besoin. Merci Papa d'avoir suivi le 13h de France 2 au commencement de tout. Merci de m'avoir littéralement coaché à chaque étape importante de ma vie aussi bien professionnelle que personnelle, merci pour tout le temps que tu as passé à relire ma thèse, et surtout merci d'être ce conseil et ce soutien indéfectibles qui me rassurent tant.

Merci Maman de toujours me rappeler de prendre soin de moi avant tout, merci de l'amour dans lequel nous baignons depuis toujours, et qui nous a permis de tant nous épanouir. Merci Sam d'être ce frère sur lequel je peux toujours compter, merci d'être une aussi grande source de fierté (et aussi de faire semblant de ne pas remarquer que je fais exprès de perdre à FIFA pour te faire plaisir). On ne choisit pas sa famille, mais si je l'avais pu, c'est vous que j'aurais choisi…

Un mot de plus pour exprimer ma gratitude à Bruno Arnaldi, dont l'appui et le soutien ont toujours été décisifs, en particulier durant ma dernière année de thèse, marquée par de nombreuses péripéties « administratives » que je n'aurai certainement pas surmontées sans lui. Merci d'avoir été à l'écoute, et finalement merci d'avoir toujours trouvé une solution à mes problèmes même quand ils paraissaient insurmontables.

Merci également à Marc Christie, d'avoir fait partie de mon CSI, et d'avoir régulièrement et de façon toujours très constructive, suivi l'évolution de mes travaux.

J'ai plaisir à mentionner et à adresser un remerciement tout particulier à Nathalie Denis, notre assistante d'équipe. J'ai toujours dit qu'un déplacement qui se passe bien, commence impérativement dans le bureau de Nathalie. Mais si ce n'était que ça ! Nathalie, tu es incontestablement l'une des raisons principales qui font que l'environnement de l'équipe est si agréable. Merci pour ton sens de l'écoute, ta gentillesse et ton efficacité.

Un grand merci à toi, Florian Nouviale, pour ton efficacité et ta compétence, ainsi que pour

tous les movies foodies organisés chez toi.

Merci également à toi Camille Jeunet, de m'avoir « tenu la main » au départ en m'aidant à réaliser mes toutes premières expériences ! de m'avoir appris tant de choses sur les BCI, de m'avoir présenté à tous ces chercheurs que je ne connaissais que de nom, lors de nos conférences ensemble. Je garderai en mémoire ton sens de la rigueur et ton approche si humaine de la recherche. Merci aussi pour tous ces moments moins formels, les « pouloulous » et les longues soirées de conf ! Je suis très heureux de te compter parmi mes amies, et je te souhaite de tout cœur la carrière exceptionnelle que tu mérites et qui se dessine déjà.

I would also like to thank Jussi Lindgren as well, for all the memorable moments we shared during our time together in Rennes. Dear Jussi, I will always remember your very unique sense of humor, and very Finnish sarcasm. . .

J'ai également un merci tout particulier et toute ma reconnaissance à adresser à Manel. Pour plein de choses en fait. Pour avoir su m'écouter et me soutenir dans les moments de doute, pour sa gentillesse, et un immense merci pour son aide dans la conception de mes slides.

Merci aux stagiaires (aujourd'hui doctorants à leur tour) avec qui j'ai eu le plaisir de travailler durant cette thèse : Jimmy Petit et Martin Guy. J'ai beaucoup apprécié de collaborer avec vous, et je ne doute pas que vos thèses seront excellentes.

J'adresse un grand merci également à Inria et l'IPL BCI-Lift, pour avoir financé mes travaux de thèse, ainsi qu'à l'équipe des BCI de France, Maureen Clerc, Fabien Lotte, Laurent Bougrain, Léa Pillette, Sébastien Rimbert, Jelena Mladenovic, Aline Roc, Camille Benaroch, Thibault Monseigne et les doctorants du Gipsa Louis Korczowski et Pedro Robrigues. Ce fut autant un plaisir de travailler avec vous que de vous retrouver en de nombreuses occasions scientifiques.

C'est un remerciement très spécial que j'adresse à l'équipe Hybrid, ma famille durant ces trois années ! Yoren Gaffary et Guillaume Cortes, mes co-bureaux, ce fut un plaisir de partager ces années avec vous. Merci également à l'ensemble des Hybrid, passés et encore présents Adrien Reuzeau, Alexandre Audinot, Andéol Evain, Antoine Costes, Antonin Bernardin, Benoit Legouis, Carl Jorgensen, Diane Dewez, Emeric Goga, Etienne Peillard, Flavien Lécuyer, François Lehéricey, Giulia Lioi, Guillaume Bataille, Guillaume Claude, Guillaume Gicquel, Guillaume Moreau, Guillaume Vailland, Gwendal Fouché, Gwendal Lemoullec, Hugo Brument, Jean-Marie Normand, Lorraine Perronnet, Mathis Fleury, Maud Marchal, Mélanie Cogné, Rebecca Fribourg, Romain Lagneau, Ronan Gaugne, Thierry Gaugry, Thomas Howard, Thomas Rinnert, Tiffany Luong, Valérie Gouranton, Victor Mercado, Vincent Goupil, Xavier De Tinguy ainsi que Théo Perrin, Charles Faure et Caroline Lebrat.

Merci également à ma deuxième équipe, les Loki ! De m'avoir toujours chaleureusement

reçu à Lille. Merci à Stéphane, Thibault, Axel, Nicole, Sylvain, Thomas, Mathieu, Ed et Amira.

Je tiens à remercier les équipes d'Inria, et tout particulièrement Anita Hennetier, Myriam Vinouze et Sophie Viaud pour leur aide à chaque fois que j'ai frappé à leur porte, et Dieu sait que je ne m'en suis pas privé.

Je remercie également les équipes de l'INSA de Rennes, mon université d'inscription. Et plus particulièrement, Aurore Gouin et Justine Gromaire pour leur gentillesse et pour leur aide dans les différentes démarches administratives que j'ai eu à accomplir.

During my thesis, I had the chance to spend three months at the Institute of Neural Engineering of TU Graz. It has been an amazing stay, from which I keep formidable memories. For it, I would like to thank you Reini, for the chance and the opportunity you gave me to learn from you, and to collaborate on an amazing project. I will value your advices and teachings for a long time.

I would also like to thank Pr. Gernot Müller-Putz for hosting me in his team, and for his help. My time in Graz has been one of the best experiences I've had and I would like to thank all of Selina, Andy, Andreea, Catarina, Lea, Petra, Aleksandra, Karina, Maria, Kathi, Joana, Jonas, Reinmar, Gudrun, Thomas and all the wonderful people I've met there.

Que toutes les personnes enfin que j'ai omis de citer nommément, parce que je dois tant à tant de personnes, sachent que je réalise que leur aide et leur soutien ont tout autant contribué à la réussite de ce travail, et que ma profonde gratitude leur est acquise.

A toutes et à tous, un immense MERCI.

TABLE OF CONTENTS

INTRODUCTION

Interacting through the power of the mind is a widespread fantasy in popular culture. Communicating through "thoughts" or manipulating virtual environments with the activity of the brain have been depicted in numerous movies and works of fiction (Figure 1).

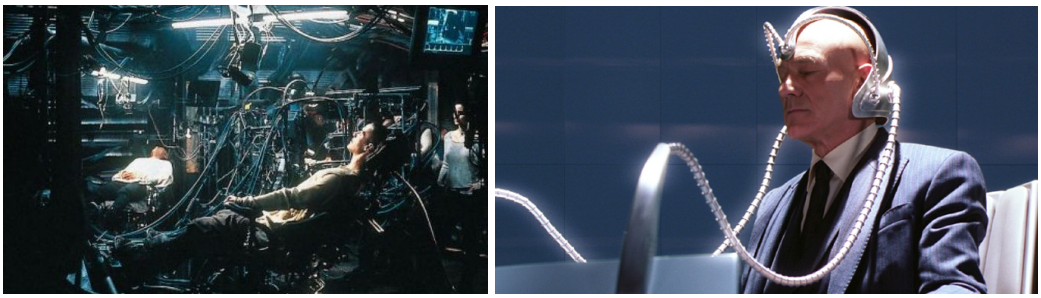


Figure 1: Examples of movies employing the concept of brain-computer interaction. In the Matrix movie (Left), the protagonists use an invasive Brain-Computer Interface (BCI) to immerse themselves in a virtual environment. In X-men (Right), one of the main protagonists uses a non-invasive BCI to sense the outside world.

The global aim of this work is to study the use of Brain-Computer Interfaces (BCIs) in Augmented Reality (AR). In particular, the underlying aspects in terms of BCI paradigms [1], new types of interfaces and system usability. BCIs hold the promise of interfacing the brain and the computer. AR aims at mixing virtual elements in the real world environment [Azuma, 1997]. Together, BCIs and AR enable hands-free interaction with virtual elements in the real world, which can be used in numerous application domains and fields such as robotics, health or entertainment. However, the combination of these two scientific domains also opens numerous research questions. Among them: What kind of brain activity can be used in AR? How can we exploit BCI outputs for interacting with real and virtual elements? How do we build usable systems combining BCI and AR?

Up to the day these lines are written, it is impossible to comprehensively decode "thoughts" and the brain activity, as traditionally understood by the general public. Instead, BCIs make use of documented neurophysiological properties and reactions of the brain to infer mental

1. A BCI paradigm is a type of brain activity or reaction that is exploited for interaction.

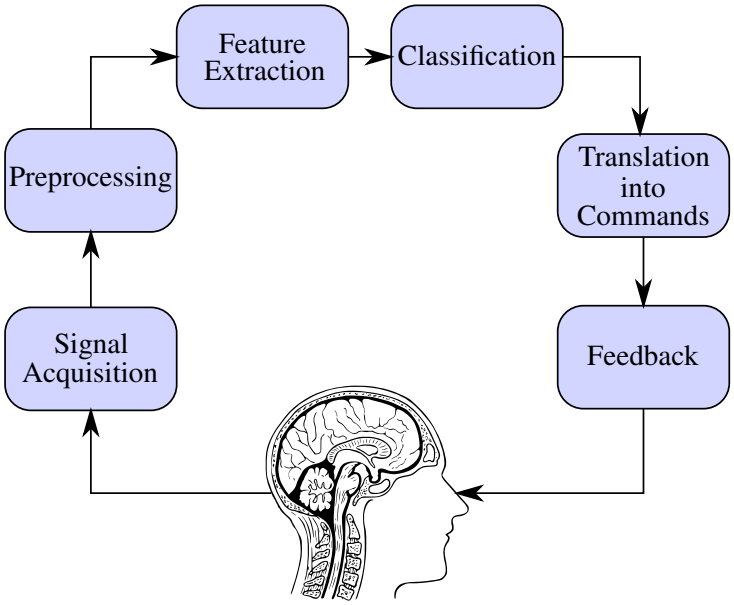


Figure 2: Illustration of the BCI processing pipeline. This pipeline, as inspired from [Mason and Birch, 2003], summarizes the different components of an online BCI.

states and desired commands. The choice of one property over others falls to the designer, considering the task the BCI is designed for. Then, to transform the activity of the brain into a computer command, a BCI follows a series of steps summarized in Figure 2. When the desired neurophysiological property has been chosen, the first step is to *measure* the brain activity. For this, several brain imaging techniques have been developed over the years. The most widely used though, and also the oldest one is called Electroencephalography (EEG) [Berger, 1929]. EEG measures the electrical activity of the brain through electrodes placed at different regions of the scalp, and gives a signal representing its temporal evolution. However, this obtained signal is always contaminated with noise. Hence, the next step called *preprocessing* consists in separating this noise from the useful signal. The following step, known as *feature extraction*, is to extract useful information out of the preprocessed signal. This step has 2 objectives: (1) reduce the dimensionality of the signal and (2) describe the signal with meaningful features. These features are then *classified* into mental states which are later translated into commands that are sent to the computer system. This whole process, also known as the "*Online BCI processing pipeline*" is often completed by a feedback step to help the user adjust the modulation of his brain activity.

On the other hand, Augmented Reality is defined as the integration of virtual objects and information in the real world in real-time [Zhou et al., 2008]. Milgram and Kishino established a continuum ranging from complete virtuality to complete reality. Between them, exist different

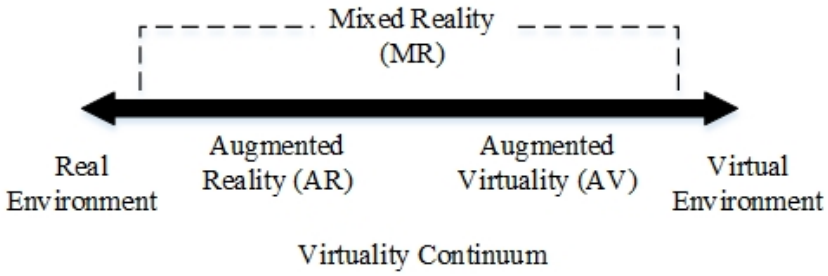


Figure 3: The "virtuality continuum" from Milgram and Kishino [Milgram and Kishino, 1994] representing the different levels of mixing between real and virtual content.

combinations of real and virtual environments [Milgram and Kishino, 1994], depending on the level of each one in the scene (see Figure 3).

According to Azuma [Azuma, 1997] three characteristics define an AR system: (1) the combination of real and virtual content, (2) the real-time interaction, (3) the 3D registration of the virtual content in the real environment. Contrarily to Virtual Reality (VR), where the user is immersed in a completely virtual world, AR mixes virtual and real content, ideally, making them appear to coexist in the same space.

Objectives

Considering the global objective of this work, of studying the combination of BCIs and AR we could identify five intermediary objectives:

1. **Introducing a new perspective of BCIs:**

   Rather than considering a BCI as a mere communication device, it is important to understand what it can offer as a source of information about the user mental state. To this day, there is no consensus on the definition of the term "Brain-Computer Interface". Several definitions have been proposed though, and they tend to reflect the diversity of the researchers' backgrounds. Medical Community may see BCIs as rehabilitation systems while a robotician may perceive them as systems to control robots. We argue that a BCI can be both. The first objective of this work is to give a novel perspective on the meaning and the terminology behind the term BCI. With this new perspective, the goal would be to better understand the integration of these technologies in AR.

2. **Studying the combination of BCI and AR in terms of BCI paradigms:**

   The second objective of this work, is to study the use of different BCI paradigms in AR. Especially because combining BCI and AR may generate interferences as EEG is very sensitive to noise [McFarland et al., 2005], it is important to assess the possibility to use every paradigm in AR. In addition, what types of brain activity is possible to detect when using a BCI in conjunction with AR? And what type of events or stimuli in AR can elicit such specific brain responses? The second objective of this work is thus to study the effects of combining BCI and AR with respect to various BCI paradigms.

3. **Improving the BCI itself:**

   One of the main limitations of current BCIs is their accuracy [Lotte et al., 2018]. In particular, one difficulty facing BCIs lies in the recognition of rest states. It is difficult to conceive systems that are able to detect the users' willingness and absence of willingness to interact. These systems, called asynchronous or self-paced, are a requirement to any practical interactive system such as in AR. Thus, the third objective of this work is to propose new methods for improving the self-paced detection accuracy of BCIs to improve the interaction in AR using a BCI.

4. **Designing new interfaces for using a BCI in AR:**

   Designing an interactive system involves two dimensions: the *performance*, i.e how well does the system work, and the *intuitiveness* of its interface, i.e how intuitive and easy to use and to understand it is. In addition to the performance of a system combining AR and BCI, our fourth objective is to characterize and evaluate the effect of the user interface (UI) in AR on the general perception of the users and the system performance.

5. **Developing new usages of BCI and AR:**

   Lastly, our fifth objective is to study the benefits of combining AR and BCI through concrete use cases.

Approach and Contributions

This manuscript describes the work and contributions that were achieved in order to meet the previously highlighted objectives. It is divided into four parts: **Part I** provides an overview of BCI technology as well as a state of the art of the combination of BCIs and AR. **Part II** describes our contributions to improve the understanding and the definition of BCIs. **Part III** describes our studies and contributions regarding the assessment of BCI paradigms that are possible to use

in AR, and **Part IV** details our contributions regarding the use of BCIs for active interaction [2] in AR. A detailed outline is given in the following:

Part I: Related Work

- **Chapter 1: Introduction to Brain-Computer Interfaces**
  The goal of the first Chapter is to provide the reader with an overview of the BCI technology. Starting from basic notions about the brain, we describe the most widely used BCI paradigms. We go through the BCI processing pipeline and present methods and algorithms used to translate brain activity into computer inputs. This chapter also provides a taxonomy of the different types of BCIs.
- **Chapter 2: Brain-Computer Interfaces and Augmented Reality**
  The goal of the second Chapter is to review the state of the art of using Brain-Computer Interfaces (BCIs) in combination with Augmented Reality (AR). First it introduces the field of AR and its main concepts. Second, it describes the various systems combining AR and BCI in the literature, categorized by their application field: medicine, robotics, home automation and brain activity visualization. Finally, it summarizes and discusses the results of the survey.

Part II: On the Definition of Brain-Computer Interfaces

In this Part, we aim at introducing a new perspective on the definition of a BCI.

- **Chapter 3: Proposing a new definition of BCI from a Human-Computer Interaction Perspective**
  In the third Chapter, we introduce a new definition of a Brain-Computer Interface. Originating from a Human-Computer Interaction point of view, we argue that a BCI should be seen as an *Interface* bridging a user through his *Brain* activity, to any *Computer* system. We discuss the previously proposed definitions, and we draw an analogy between a BCI and a mouse, as a widespread Human-Computer Interface.

2. BCI for active interaction is used here to designate scenarios where the BCI is used to actively send commands.

Part III: BCI Paradigms in AR

This Part gathers our contributions regarding the testing of different BCI paradigms in AR. Our goal is to assess the possibility to (1) use the Steady-State Visual Evoked Potentials (SSVEP) paradigm for direct interaction in AR, and (2) elicit and detect Error Related Potentials (ErrPs) in AR.

- **Chapter 4: Using SSVEP for AR-BCI interaction**
  In the fourth Chapter, we study the use of the SSVEP paradigm in AR. We investigate the effect of the AR context on the possibility to elicit and detect SSVEP responses. Then, as AR is rarely static, often involving scenarios where the user is free to move, we evaluate the effect of head motion on the SSVEP responses in AR.
- **Chapter 5: Towards the Detection of System Errors in AR Using ErrPs**
  In the fifth Chapter, we study the possibility of eliciting and detecting Error Related Potentials (ErrPs) in AR. In particular, we study the effect of common errors in AR systems (tracking errors, feedback errors and background anomalies) on users' EEG signals and the possibility to automatically detect elicited ErrPs.

Part IV: Using SSVEP-based BCI for Interacting in AR

This Part gathers our contributions regarding the use of the SSVEP paradigm for interacting in AR. Following the testing of two BCI paradigms (SSVEP and ErrP) in AR (Part III), we have chosen to focus our efforts on the SSVEP paradigm, as it is one of the most promising EEG-based BCI paradigms [Chen et al., 2015]. Our approach consisted in 3 steps: (1) improving the performance of the SSVEP-based BCI itself, (2) improving the intuitiveness of the SSVEP user interface in AR, and (3) improving the application using an SSVEP-based BCI in AR.

- **Chapter 6: Improving Asynchronous SSVEP Classification Using CCA and Hidden Markov Models**
  In the sixth Chapter, we aim at improving the self-paced classification accuracy of SSVEP responses to improve the interaction using SSVEP in AR. For this matter, we introduce *HCCA* (Hidden Markov Model based Classification of CCA Coefficients) a new method based on the analysis of the dynamics of Canonical Correlation Coefficients through Hidden Markov Models. The method is presented in details, and the results compared to the state of the art, show that HCCA outperforms existing methods in terms of classifying

no-control states, while displaying high command classification rates.

– **Chapter 7: AR Interfaces for SSVEP-based BCIs**
In the seventh Chapter, we characterize the design space of the display strategies of SSVEP user interfaces in AR. We evaluate the effect of these display strategies in terms of user perception and SSVEP classification accuracy, aiming to extract guidelines on how to tradeoff between the system's performance and intuitiveness.
– **Chapter 8: Combining BCI and AR for Home Automation**
In the eighth Chapter, we propose general guidelines towards the development of AR and BCI-based home automation systems. We propose global specifications that such systems should meet and design a generic architecture to help the development of maintainable systems. We also describe an industrial proof of concept of integrating a BCI and AR system to a commercially available home automation platform.

PART I

Related Work

CHAPTER 1

INTRODUCTION TO BRAIN-COMPUTER INTERFACES

Abstract:

This chapter aims at providing the reader with an overview of the BCI technology. Starting from basic notions about the brain, it describes the most widely used BCI paradigms. It goes through the BCI processing pipeline and present methods and algorithms used to translate brain activity into computer inputs. This chapter also provides a taxonomy of the different types of BCIs.

1.1 Introduction

The history of BCIs starts in the 19th century. The first evidence regarding the electrical nature of the brain activity was found by Richard Caton in 1872 [Caton, 1875], who observed it on animals. He implanted a galvanometer electrode in the brain of rabbits and monkeys and detected weak electrical activity. Based on this, in 1924 [1] Hans Berger performed the first recordings of Human electrical brain activity [Berger, 1929]. Thanks to a device he invented: the electroencephalogram (EEG), he was able to record the electrical activity of the brain in humans, from the surface of their scalp. The invention of EEG can be considered as the cornerstone of the research in BCI as today, it remains the most widely used technique for measuring brain activity [Nicolas-Alonso and Gomez-Gil, 2012].

However, the term ”Brain-Computer Interface” itself, only appeared in the 70s. It was introduced by Jaques Vidal [Vidal, 1973] who aimed at exploiting the measured brain activity to interact with a computer. In his paper "*BCI challenge: Control of objects using EEG signals*" he described his experiments to use EEG signals to drive a cursor on a computer screen.

Starting from there, BCIs have been an ever growing research topic [Wolpaw and Wolpaw,

1. The research was conducted in 1924 but only published in 1929

2012]. With improved understanding of the brain, advanced technological capabilities, new types and new applications of BCIs were successfully developed.

1.2 Notions of Neuroscience

BCIs extract information from the users' brain activity which is part of the nervous system. In this section, we introduce some fundamental concepts about the brain, and the origin of the brain activity.

In a cellular level, the nervous system can be divided into 2 wide categories of cells: *Neurons* and *Neuroglia*. However, the ones that are principally capable of electrical signaling and conveying information are neurons [Purves et al., 2004]. In fact, neurons can be considered as the elementary components of information processing in the nervous system.

1.2.1 The Neurons

A neuron is a cell that transmits information under the form of electrical and chemical signals over distances [Purves et al., 2004]. It is constituted of a cell body and an *axon* (a unique extension from the neuronal cell body, specialized in the conduction of signal between neurons), and ends in a *synaptic* area [Purves et al., 2004, Kandel and Tollet, 2016] (see Figure 1.1).

There are two types of electrical potentials for the signal conveyed by the neurons. The first corresponds to a depolarization of the axonal membrane, called the *action potential* or *spike*. It is a self-regenerating electrical signal that propagates through the neural body, and is conveyed by the synapses [Kandel and Tollet, 2016, Purves et al., 2004]. As for the majority of the neurons, there is no physical continuity between pre- and post-synaptic elements, they communicate through the secretion of molecules producing a *post-synaptic potential*, in an extracellular space between them called the synaptic cleft [Purves et al., 2004] as stated by Ramon y Cajal [y Cajal, 1954].

Neurons never function in isolation [Purves et al., 2004]. Assembled, they form *Neural Circuits* that are themselves assembled to form *Neural Systems*, present in the brain. These neural systems are mainly divided in 3 functional categories [Purves et al., 2004]:

1. Sensory systems that manage information about the external and internal state of the organism.
2. Motor systems that generate actions.

2. http://biomedicalengineering.yolasite.com/neurons.php

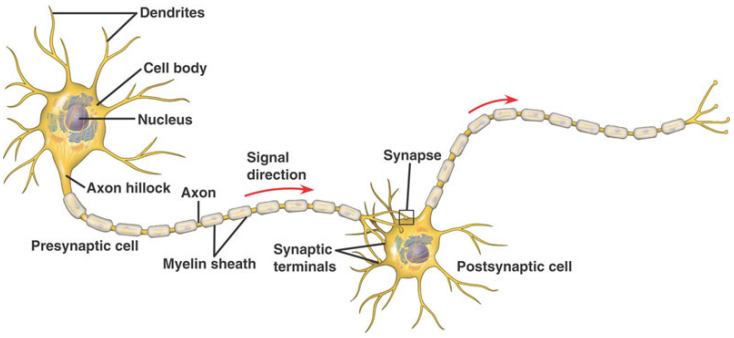


Figure 1.1: Architecture of Neurons. It is constituted of a cell body and an axon that ends in synapses [2]. Action potentials are self regenerative potentials propagating through the axon. Between the synapses of a pre-synaptic (emitting neuron) and a post-synaptic (receiving neuron) cells, a postsynaptic potential is generated by the release of neurotransmitters [Purves et al., 2004].

3. Associational systems that associate sensory and motor sides for generating higher-level functions (e.g. Attention, Emotion etc.).

1.2.2 The Brain

The Encephalon or *Brain*, is the part of the Central Nervous System that is enclosed in the cranial box [Purves et al., 2004]. It comprises the *Cerebrum*, the *Cerebellum* and the *Brainstem*. For BCI applications, we focus on the functional anatomy of the cerebrum, in which most of the areas exploited for BCI lie.

The Cerebrum is composed of two hemispheres (right and left). There are mainly 4 external lobes (see Figure 1.2): the *Frontal*, the *Parietal*, the *Occipital*, and the *Temporal* lobes [Kandel and Tollet, 2016]. These lobes can be described as specific areas of the *Cortex* (the surface of the Cerebrum) responsible of certain groups of tasks. For example, the frontal and temporal areas of the dominant hemisphere is where language is processed, whereas the center of spatial orientation is in the right parietal lobe and the organization of complex gesture is done in the frontal lobe [Kandel and Tollet, 2016].

The separation between the lobes is based on the folded aspect of the brain, resulting in bumps and fissures known as *gyri* and *sulci* resp. [Kandel and Tollet, 2016] (See figure 1.3).

3. https://en.wikipedia.org/wiki/Parietal_lobe#/media/File:Lobes_of_the_brain_NL.svg
4. http://www.qbi.uq.edu.au/the-brain/anatomy/brain-lobes

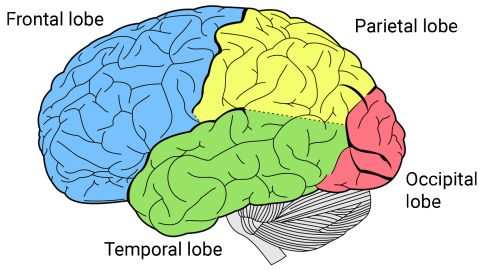


Figure 1.2: Illustration of the 4 external lobes in a human brain [3].

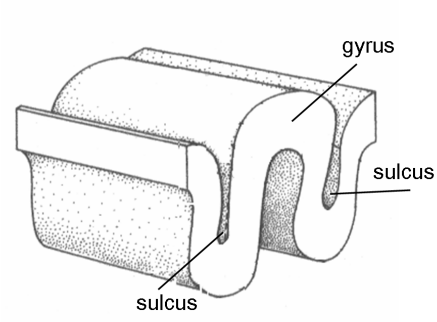


Figure 1.3: Illustration of the folded aspect of the brain. The bumps are known as the *gyri* and the fissures as *sulci* [4].

Frontal Lobe

The frontal lobe is the most *rostral* (see Figure 1.4) area of the brain. In the dominant hemisphere, it comprises the Broca's area that is considered as the main area for speech production [Broca, 1861], and the pre-motor area responsible of the preparation and planning of movement. It also makes it possible to decide which movement is needed to carry out a desired action [Kandel and Tollet, 2016]. The frontal lobe also contains the primary motor cortex [Kandel and Tollet, 2016]. Primary areas comprise the areas responsible for motor functions and sensory reception. Secondary or associative areas are responsible for more elaborate information processing and plurimodal information processing [Kandel and Tollet, 2016]. In fact, to every point of the precentral gyrus, corresponds a part of the body that it controls. This is called *somatotopy* [Kandel and Tollet, 2016] and is commonly illustrated by a map known as the *cortical homunculus* (See Figure 1.5).

5. http://www.corpshumain.ca/en/Cerveau1_en.php

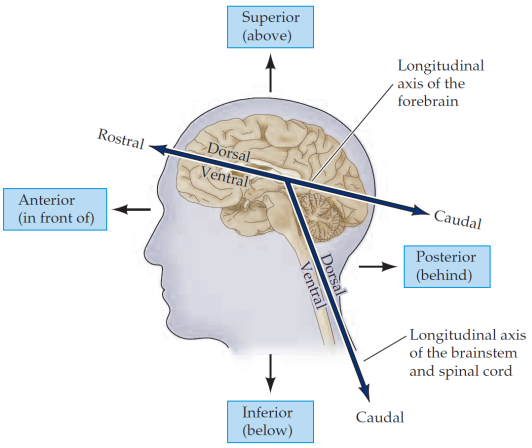


Figure 1.4: Illustration of the anatomical terminology relative to the positions of brain areas (image from [Purves et al., 2004]).

Parietal Lobe

The parietal lobe houses the primary somatosensory cortex which receives and interpret sensory information (e.g. pain, temperature, touch...) [Kandel and Tollet, 2016].

As for the motor homunculus, the organization of the primary somatosensory cortex gives rise to a sensory homunculus [Kandel and Tollet, 2016] (see figure 1.5). The majority of sensory information comes to the parietal lobe through the *thalamus*, which is involved in the regulation of consciousness, sleep and alertness in addition to the relaying of motor and sensory information to the cerebral cortex [Sherman and Guillery, 2001].

Occipital Lobe

The occipital lobe is the most *caudal* of the lobes, it is separated from the parietal lobe by the occipito-parietal sulcus and from the temporal lobe by the pre-occipital notch [Kandel and Tollet, 2016]. It hosts the primary visual cortex [Purves et al., 2004]. Once again, as for the primary motor cortex and the primary somatosensory cortex, there is a systematic correspondence between every point in the visual field and a specific area in the primary visual cortex. This property is known as the *retinotopy* [Kandel and Tollet, 2016].

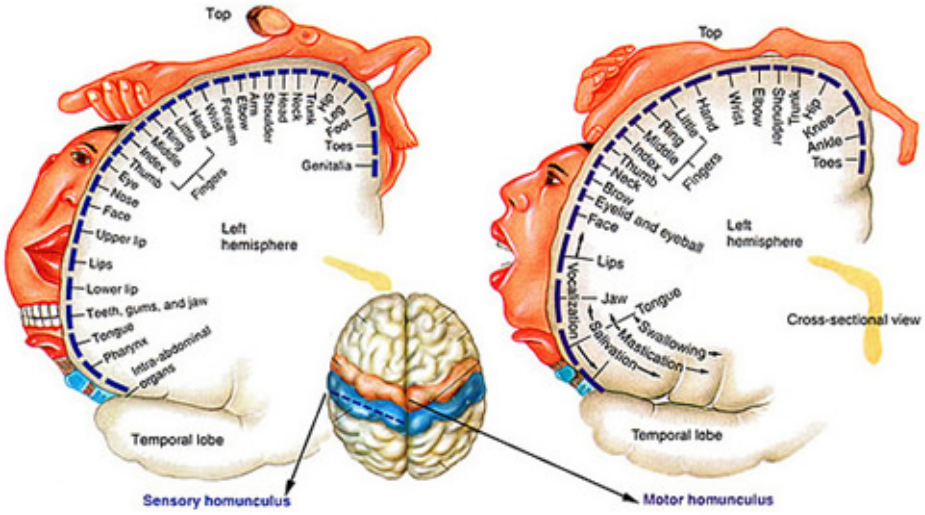


Figure 1.5: Illustration of the sensory (Left) and motor (Right) Homunculi [5].

Temporal Lobe

The temporal lobe is located in the lower face of the cerebrum. It is composed of the primary auditory cortex, the associative auditory cortex and the associative temporal cortex involved in language memory [Kandel and Tollet, 2016].

The information coming from each auditive nerve is bilaterally projected in the primary auditory cortices, and it is therefore possible to define a *tonotopy* (activated areas associated with specific sound frequency) [Kandel and Tollet, 2016].

1.3 Measurement of Brain Activity

A BCI works by associating a mental state to a computer command. These mental states are usually determined through the observation of changes in brain properties at specific regions. BCIs based on motor tasks will focus on changes in the motor cortex, while BCIs based on visual properties will focus on the occipital lobe. In order to monitor these changes, several tools allow to measure the activity of the brain [Weiskopf et al., 2004]. The main difference between these tools lies in the physical property of the brain that they sense. We mainly distinguish 2 groups of measurement tools:

The first one measures the metabolic changes in the cerebrum, particularly the changes in the blood flows and the changes in levels of hemoglobin through the brain. As the brain regions that

Figure 1.6: Examples of brain activity measurement devices namely: an EEG (top left), an fNIRS (top right), an MEG (bottom left) and an fMRI (bottom right) [6].

operates need energy, meaning oxygen to function, these techniques are based on the "*Blood Oxygen Level Dependent* (BOLD) effect" [Ogawa et al., 1990]. Among these techniques, we can cite *functional Magnetic Resonance Imaging* (fMRI) and *functional Near-Infrared Spectroscopy* (fNIRS). While fMRI measures the levels of hemoglobin through the magnetic properties of blood, fNIRS measures it through the differences in the level of absorption of near-infrared lights.

6. Image Sources
Top Left:
`https://www.brainlatam.com/manufacturers/sync-lab/active-and-passive-eeg-electrode\-integration-221`
Top Right:
`https://en.wikipedia.org/wiki/Functional_near-infrared_spectroscopy#/media/File:`

The second one measures the electrical activity generated by populations of neurons and the underlying differences of potentials in different regions of the cortex [Vidal, 2016]: *Electroencephalography* (EEG) and *magnetoencephalography* (MEG). While EEG directly measures the electrical activity of the neurons (as it propagates to the scalp), MEG measures the differences in magnetic fields generated by the electrical activity of the neurons.

Amongst the main characteristics of measurement techniques are the *temporal* and *spatial* resolutions. Electrophysiological techniques holds, in general, better temporal resolution, meaning that a change in the brain state will quickly translate in a change in the measurement, while due to the latency of changes in blood flow, techniques measuring the level of oxygenation have poorer temporal resolution [Perronnet et al., 2017]. On the other hand the latter have a better spatial resolution, as they measure the differences in blood oxygenation in 3D while electrophysiological measuring techniques only observe a projection and mixture of the source signal on the scalp [Perronnet et al., 2017].

The common point between the previously cited measurement techniques though, is the fact that they are all based on external measurement. They measure brain activity from outside the surface of the head. They are called *non-invasive*. Other techniques exist that measure the brain activity from inside the skull and require a surgery to implant sensors. They are called *invasive* techniques. For example, *electrocorticography* (EcoG) measure the differences of potentials directly on the surface of the brain, using electrodes implanted in the skull [Wolpaw et al., 2006]. Considering the risks inherent to surgeries, invasive techniques are not suitable as interaction media for healthy users and their use is limited to particular medical cases. In the scope of this thesis, we focus on EEG, as it is the most widely used measurement technique for BCIs [Nicolas-Alonso and Gomez-Gil, 2012], especially when developing interactive BCI for healthy users. There are at least 2 main reasons for that: First, EEG has an excellent temporal resolution, in the order of millisecond (ms), which makes it the best measuring technique to detect quick responses in the brain activity or for real-time interaction. The second reason is that EEG has a relatively lower cost when compared to other techniques. Depending on their precision, sampling rate and other parameters, medical EEG systems costs, range between 8000$ [7] and 30000$ [8] for the

FNIRS_head_Hitachi_ETG4000_6.jpg
Bottom Left:
https://en.wikipedia.org/wiki/Magnetoencephalography#/media/File:NIMH_MEG.jpg
Bottom Right:
https://www.ndcn.ox.ac.uk/divisions/fmrib/what-is-fmri/introduction-to-fmri

7. https://mbraintrain.com/
8. http://www.gtec.at/

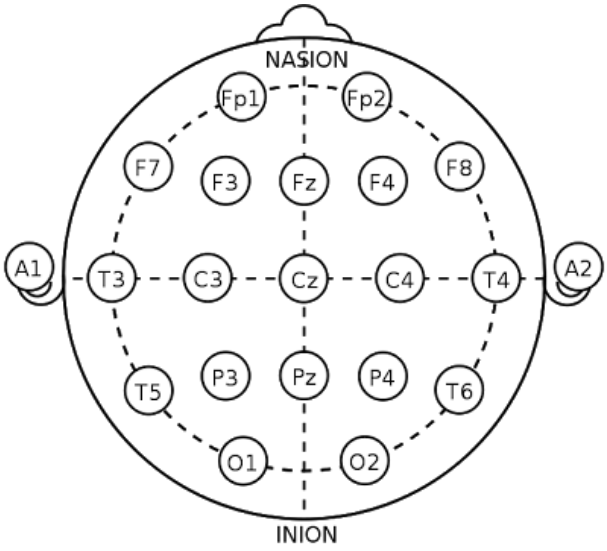


Figure 1.7: Illustration of the 10-20 system for EEG electrode placement over the scalp [10].

medical systems and can be as cheap as 200$ [9] for the general public.

1.3.1 Electroencephalography

Electroencephalographs usually appear under the form of headsets or caps equipped with electrodes that are placed on the surface of the scalp. These electrodes are often placed following a standard mapping known as the 10-20 system [Jasper, 1958] (see Figure 1.7). In order for the signal to be exploitable, the electrodes are connected to an amplifier that amplifies and digitalizes the recorded signal. In order to lower the impedance between the electrodes and the skin and thus, to improve the signal quality, electrodes may require an electrolyte: conductive gel or saline solutions. However, some alternatives exist using dry electrodes, i.e which do not require electrolytes.

For its most part, the electrical activity measured by EEG is generated by the postsynaptic potentials of large populations of neurons [Vidal, 2016]. Each electrode measures the difference of electrical potential between its location and a reference electrode. The evolution of the amplitude over time is then represented as a temporal signal (see Figure 1.8) which indicates the brain activity of the BCI user.

EEG signals can usually be decomposed in components or oscillations called ”rhythms” [Nie-

9. https://choosemuse.com/muse/
10. https://en.wikipedia.org/wiki/10-20_system_(EEG)

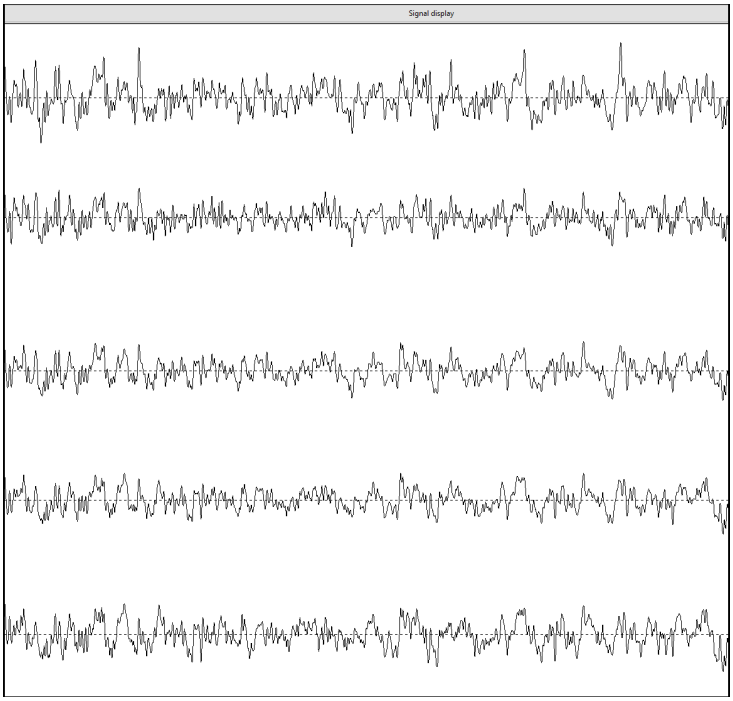


Figure 1.8: Epoch of EEG signal recorded from 5 electrodes.

dermeyer, 2005]. These rhythms are characterized by their spatial and spectral localizations and are often used as indicators of changes in mental state. Six brain rhythms have been identified:

- **Delta (δ) rhythm:** Slow oscillations located at a frequency band between 1 and 4Hz. Delta oscillations are found during adults deep sleep [Maquet et al., 1997].
- **Theta (θ) rhythm:** Oscillations located at a frequency band 4 and 7Hz. Theta oscillations are found during drowsiness or in young adults [Cantero et al., 2003].
- **Alpha (α) rhythm:** Oscillations located at a frequency band between 8 and 13Hz. Alpha oscillations are mainly found in the posterior regions of the brain in subjects having closed eyes, or in relaxation states. This was the first brain rhythm observed by Hans Berger [Berger, 1929].
- **Mu (μ) rhythm:** Oscillations at the same frequency band as Alpha, between 8 and 13Hz. However, Mu oscillations are found in the motor and sensorimotor cortex. The amplitude of this oscillations changes when the subject performs a movement. Thus, this brain rhythm is also known as the "*Sensorimotor rhythm*" [Kozelka and Pedley, 1990].
- **Beta (β) rhythm:** Oscillations at a frequency band between 13 and 30Hz. Beta oscillations are observed in awakened and conscious states. The amplitude of this rhythm changes

when the subjects perform movements [Kozelka and Pedley, 1990].

– **Gamma (γ) rhythm:** Oscillations at frequencies higher than 30Hz. This rhythm is associated with various cognitive and motor functions [Vanderwolf, 2000, Jensen et al., 2007]. For example, high gamma activity at temporal locations is associated with memory processes [Malik and Amin, 2017]

1.4 Types of BCI

There are several criteria under which it is possible to classify BCIs. In addition to the measurement technique, it is also possible to classify a BCI depending on whether or not it requires residual muscle activity. *Independent* BCIs can be operated without any muscle activity while *dependent* BCIs demands some muscular control. For example, a BCI that would require the subject to overtly gaze at some stimulus is a dependent BCI [Allison et al., 2007].

Another criterion to classify a BCI is based on the nature of the interaction they provide the users with. *Synchronous* BCIs require the users to perform mental tasks, i.e send a command at specific moments determined by the system [Pfurtscheller et al., 2006]. These systems are easier to implement as it is possible to know the precise time window where to search for a brain pattern. *Asynchronous* BCIs on the other hand, also known as *Self-Paced*, allow the user to interact at will. They are closer to the definition of interactive systems [11] because the user can choose to interact or not at unpredictable moments. However, asynchronous BCI are more difficult to implement as they require the continuous analysis of brain activity as they do not hold any information on the time when the mental task is being performed.

Finally, it is possible to classify BCIs by considering the nature of the neurophysiological properties they exploit for the interaction. Also called *BCI Paradigm*. It is possible to identify 3 types of BCIs depending on the kind of paradigm they exploit (*Active*, *Reactive* and *Passive*) [Zander and Kothe, 2011]:

1.4.1 Active BCI

Active BCIs are the ones where users have to voluntarily and spontaneously modulate their brain activity without any external stimulation, as schematized in Figure 1.9. The modulation is done through the performing of a mental task that the system has to be able to detect and

11. Interactive system is a system depending on the inputs originating from another system it does not control [Goldin et al., 2006]

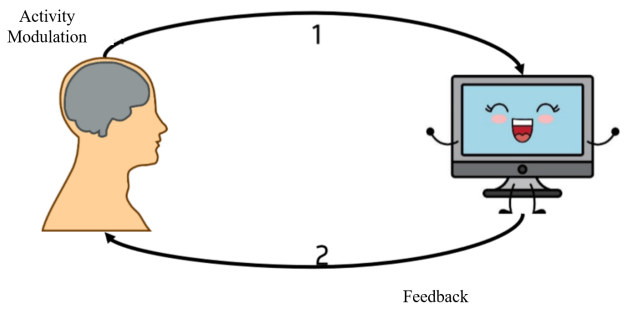


Figure 1.9: Schematization of the functioning of an active BCI. The users spontaneously modulate their brain activity to (1) send commands to a computer system, which provides feedback (2) so that users can adjust their modulation.

associate with a command. The detection of the realization of the mental task is done through the monitoring of neuromarkers [12] associated with the mental tasks. Most of the time, the neuromarkers used to drive active BCIs are the amplitude of motor and sensorimotor rhythms [Yuan and He, 2014]. These types of neuromarkers usually referred to as *Spontaneous Signals*, among them we can cite:

Mental Imagery

Mental imagery regroups the situations where an individual is supposed to imagine performing a mental task [Jeunet et al., 2016b] (limb movement imagination, mental arithmetic, generation of words etc.). These mental tasks are associated with changes in the brain signals at specific frequency bands and specific locations [Benaroch et al., 2019].

For example, motor imagery is a BCI paradigm based on the property that when individuals imagine a movement of their limbs (right/left arm or feet for example), a contralateral (in the hemisphere at the opposite side of the limb) decrease of power (Event Related Desynchronization ERD) in the μ and β bands is observed over the motor region of the cortex [Pfurtscheller and Neuper, 2001] (see Figure 1.10). By associating a command to each motor imagery task, it becomes possible to drive a multiple-commands system.

12. Neurophysiological features associated with mental states.

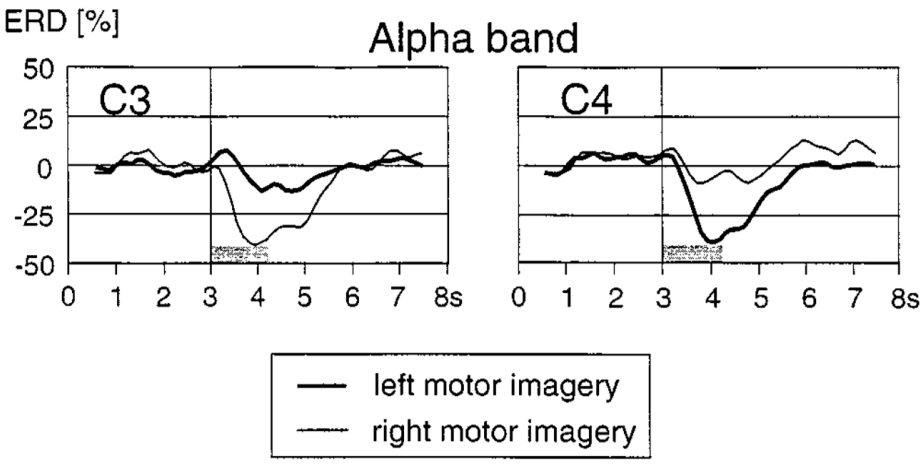


Figure 1.10: Event Related Desynchronisation (ERD) associated with left and right arm imaginary movement. The movement imagination starts at 3s. (Picture taken from [Pfurtscheller et al., 2000]).

Slow Cortical Potentials

Slow cortical potentials are slow variations of the cortical activity (average cortical potential). It has been shown that a user is capable to make these variations positive or negative using "*operant conditioning*". Through very long sessions of training, a user is capable to learn how to voluntarily change the amplitude of his/her sensorimotor rhythms [Wolpaw and McFarland, 2004, Wolpaw et al., 1991, Vaughan et al., 2006, Wolpaw, 2007] (see Figure 1.11). During operant conditioning, the users are free to choose whatever mental strategy they find the most efficient [Birbaumer et al., 2003]. Thanks to a given feedback, the users can learn what to change in their mental strategy to reach the intended goal. By setting up a threshold on the amplitude of the cortical activity, it is possible to drive a binary command system.

Covert Visuo-Spatial Attention

There are 2 types of visual attention: *Overt* and *Covert* [Kulke et al., 2016]. Overt visual attention is the fact of focusing on a specific direction by shifting the gaze towards the subject of attention. Covert attention on the other hand, corresponds to a shift of attention towards a particular direction without any eye movement.

Studies have shown that covert attention is associated with an increase in the α power in

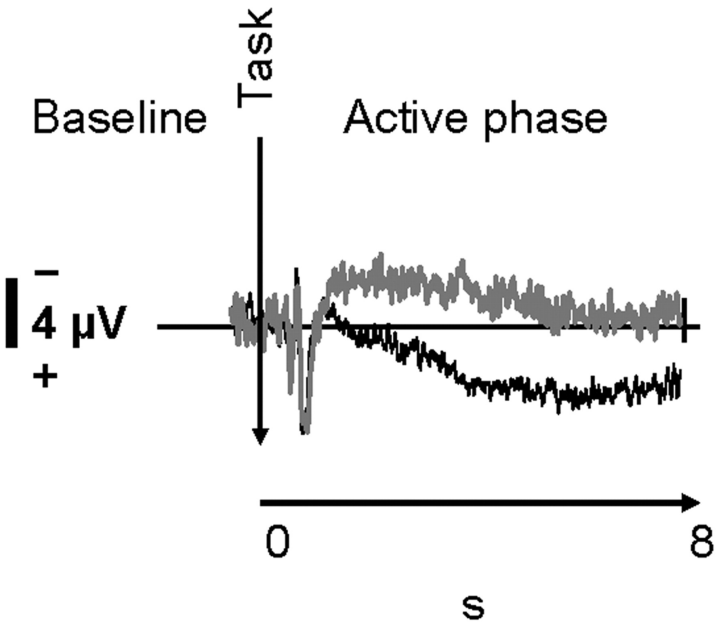


Figure 1.11: Illustration of the evolution of Slow Cortical Potentials. (Picture taken from [Strehl et al., 2006]).

the ipsilateral (same hemisphere) occipital lobe, and a decrease in the contralateral (opposite hemisphere) occipital lobe [Tonin et al., 2013, Jeunet et al., 2017]. In other words, when an individual covertly focuses on the right (resp. left) side of their visual field, the power of the α band on the right (resp. left) side of the occipital area increases, while it decreases in the left (resp. right) side of the occipital lobe.

By monitoring the difference in the α level between the right and left sides of the visual cortex, it is possible to determine on which side the user's covert attention is focused, and drive a binary command system [Tonin et al., 2013].

1.4.2 Reactive BCI

Reactive BCIs allow users to voluntarily interact and send commands to a computer [Zander and Kothe, 2011]. However, unlike active BCIs, the brain activity is not spontaneously modulated. Instead, reactive BCIs exploit the reactions of the brain to external events or stimuli on which the user has to focus his/her attention as schematized in Figure 1.12. The neuromarkers used to drive this type of BCI are referred to as *Evoked Potentials (EP)* [Cabestaing and Derambure, 2016].

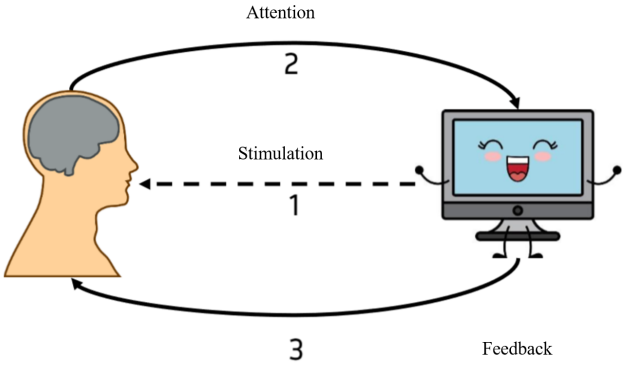


Figure 1.12: Schematization of the functioning of a reactive BCI. The system provides the users with sensory stimulations (1). Based on the stimulation on which they focus their attention (2), a command is determined and a feedback (3) is provided.

P300

When an expected and rare event occurs, a particular signal, characterized by a positive peak approximately 300ms after the occurrence of the event can be observed in the parietal and occipital parts of the brain [Chapman and Bragdon, 1964] (see Figure 1.13 Left).

The P300 potential is usually triggered using the oddball paradigm [Polich and Margala, 1997], where the user is subject to a random sequence of flickering stimuli. If the object attended by the user flashes, a P300 is elicited in the EEG signal.

The typical use case of the P300 is the P300-speller [Farwell and Donchin, 1988]. It displays to the users, a grid of letters randomly highlighted. The users are then asked to focus on and count the number of times the letter they want to spell is highlighted (see Figure 1.13 Right). When this intended letter flickers, a P300 appears. When detecting a P300, it is possible to determine which letter the user selected by determining the letter that was highlighted 300ms earlier.

Steady State Evoked Potentials

In general, an evoked potential is a brain response to an external sensory stimulation [Cabestaing and Derambure, 2016]. Depending on the nature of the stimulation, the evoked potential appears on the corresponding brain area. By periodically repeating the stimulation, neurons synchronize their firing with the repeated stimulus and that generates a *Steady-State Evoked potential (SSEP)* [Regan, 1977].

When the stimulus is visual, the potential can be observed over the occipital area and is

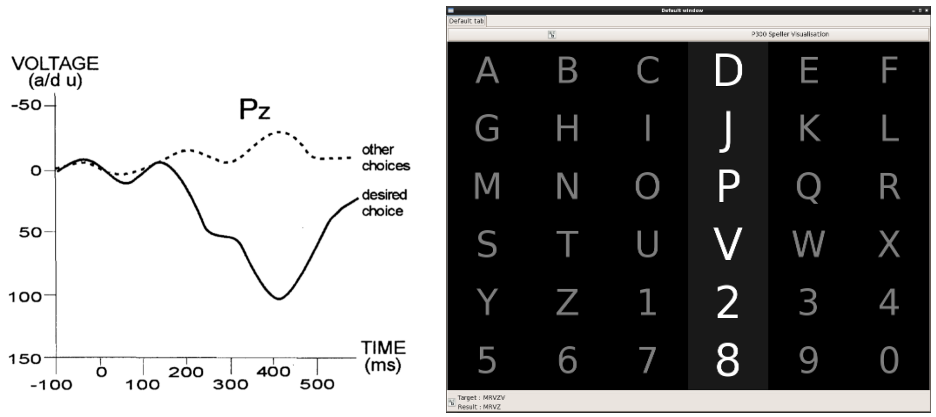


Figure 1.13: (Left) Illustration of the P300 potential (Picture taken from [Wolpaw et al., 2002]). (Right) Illustrative application of the P300, the *P300 speller*. Rows and columns blinks randomly while the user is focusing on one letter. Every time the row or column of the intended letter blinks, a P300 can be expected.

called *Steady-State Visual Evoked Potentials (SSVEP)* [Regan, 1989]. When the stimulus is auditory, the steady-state potentials, called *Steady-State Auditory Evoked Potential (SSAEP)* can be extracted from the auditory cortex at the temporal lobe, while tactile stimulations generate *Steady-State Somatosensory Evoked Potential (SSSEP)* which can be extracted from sensory area in the pre-frontal cortex.

The common element between these potentials is that they manifest under the form of an increase of the power of the brain signal at the same frequency of the stimulation and its harmonics (see Figure 1.14).

1.4.3 Passive BCI

The original purpose of a BCI was to control a computer or a device. It is however also possible to use a BCI for monitoring the general mental state of the users [Roy and Frey, 2016]. In Passive BCIs, users are not intended to voluntarily control the system [Zander and Kothe, 2011]. Instead, their brain activity and mental state is passively monitored so that the system content or behaviour can be eventually adapted accordingly [Yuksel et al., 2016] (see Figure 1.15).

Cognitive Workload

There is no consensual formal definition of cognitive workload. In general, the term is used to designate the mental effort required from the user to accomplish a task with variant levels

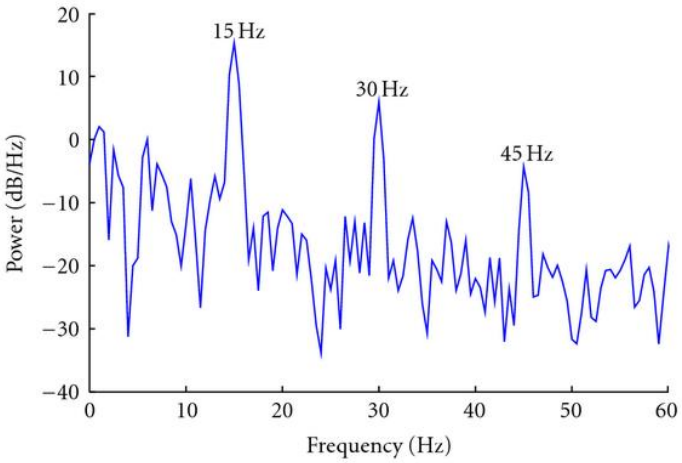


Figure 1.14: Illustration of the Power Spectral Density (PSD) of an EEG signal, where the user is focusing on a visual stimulation at 15Hz. It is possible to observe peaks at the stimulation frequency and its harmonics (Picture taken from [Zhu et al., 2010]).

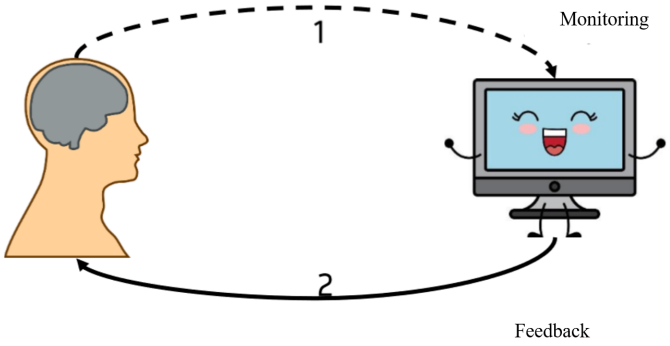


Figure 1.15: Schematization of the functioning of a passive BCI. Users are not intended to voluntarily control the computer system. Their brain activity is being passively monitored (1) and the computer system may use it to adapt its content or send feedback to the user (2).

of difficulty [Gevins and Smith, 2006]. A very popular task to modulate the level of cognitive load is the N-Back task [Herff et al., 2014]. It consists in displaying a sequence of letters to the users, asking them to perform an action (typically pressing a key) when the displayed letter is the same as the one displayed N steps before [Evain et al., 2017]. Naturally, increasing the number N increases the working memory demand, and thus the cognitive load of the user.

In terms of neuromarkers, previous studies [Gevins and Smith, 2000, Missonnier et al., 2006, Antonenko et al., 2010] have shown that the increase in cognitive load was associated with a decrease of power in the α band and an increase of power in both θ and δ bands over the parieto-central area. By determining the powers of these frequency bands, it is possible to estimate the level of Cognitive Workload of the user.

Typical scenarios of real-time estimation of the cognitive load consist in adapting the interactive system content to the level of cognitive load of the user, either by increasing the difficulty of a task to keep him/her in a state of flow [13] [Csikszentmihalyi, 1975, Blankertz et al., 2007, Yuksel et al., 2016] or reducing the amount of displayed content when the user in high workload state [Herff et al., 2014].

Error Related Potentials

When individuals notice the occurrence of what they perceive as an error, a specific signal pattern can be observed from the EEG data. This potential, referred to as *"Error-Related Negativity"* (ERN) was first highlighted in the early 90s [Falkenstein et al., 1991, Gehring et al., 1993], is characterized by a negative deflection over the fronto-central area, around 50 to 100ms after the error onset, followed by a positive peak in the centro-parietal area of the scalp [Falkenstein et al., 2000] (see Figure 1.16).

Several studies have shown that the amplitude of the ERN was modulated by the importance of the error with regards to the given task [Frank et al., 2005], as well as with the subjective awareness of the error [Falkenstein et al., 2000, Navarro-Cebrian et al., 2013, Wessel, 2012]. In addition to the ERN, a similar EEG pattern was shown to appear after an erroneous feedback. Called the *"Feedback-Related Negativity"* (FRN), this potential appears between 200 and 300 ms after the feedback onset [Hajcak et al., 2005]. Today, these patterns are grouped under the term *"Error-Related Potentials"* (ErrP), which has become widespread notably in the HCI community [Chavarriaga et al., 2014]. Up to when these lines are written, at least 4 different types of ErrPs have been identified [Spüler and Niethammer, 2015a], depending on the circumstances in which

13. The state of flow is the mental state where a person feels when completely immersed and focused on a task, not too difficult and yet challenging.

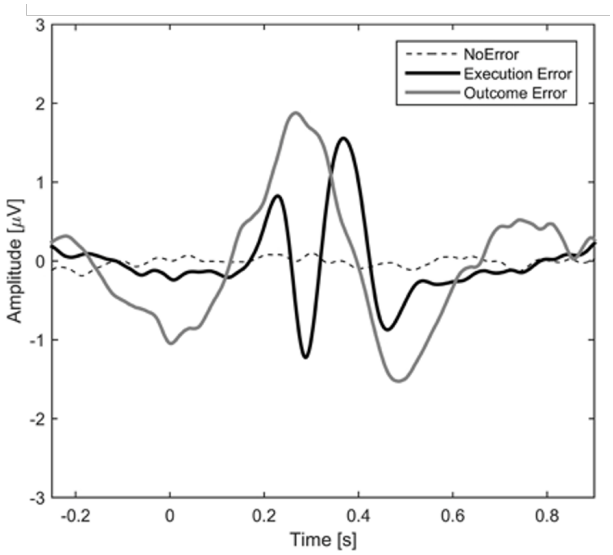


Figure 1.16: Illustration of different Error Related Potentials (picture taken from [Spüler and Niethammer, 2015a]).

the error eliciting them occur:

1. *Execution errors*: correspond to errors made by the user during the execution of a task [Milekovic et al., 2012].
2. *Outcome error*: correspond to the situation where the outcome of an action differs from what was expected [Milekovic et al., 2012].
3. *Observation errors*: correspond to the observation of another person or agent committing an error [van Schie et al., 2004].
4. *Interaction errors*: correspond to the execution of an unintended action by the interactive system [Dias et al., 2017].

Once the BCI paradigm to use has been chosen, and the subject's brain activity has been recorded, the analysis of the signal follows the BCI processing pipeline.

1.5 BCI Processing Pipeline

In this section, our goal is to present some of the most employed tools for analyzing EEG signals following the steps of the BCI processing pipeline (preprocessing, feature extraction and classification).

There are 4 different possible representations of EEG signal [Clerc, 2016]: *temporal*, *spectral*, *spatial* and *statistical*. The difference between them lies in the type of information they emphasize relatively to the paradigm being designed.

1.5.1 EEG Representation

Temporal Representation

The most basic way to represent EEG signal is under the form of a set of time series. For each electrode, the amplifier provides a time signal obtained through sampling and quantification, that represents the evolution of the amplitude recorded by the electrode after each sampling interval. With this representation, EEG signal is usually decomposed into *epochs* that represent the elementary unit that is analyzed. For example, a 2s-long epoch, with a sampling frequency of 512 Hz is represented by a vector of 1024 points for each electrode channel.

Most of the time, successive EEG epochs are constructed to overlap, meaning that the new epoch starts before the previous one ends [Clerc, 2016]. In addition to augmenting the quantity of available data, the idea behind this mechanism of overlapping epochs is to make sure to capture of the onset of the mental task as when they overlap, chances are higher that one epoch will start with the start of the mental task, especially in the context of asynchronous classification.

Spectral Representation

Since their very first recordings, it appeared that EEG signals had an oscillatory character [Clerc, 2016]. EEG signals can be seen as a sum of sub-signals of different frequencies originating from different regions of the brain. Thanks to the Fourier decomposition [baron Fourier, 1822], every temporal signal can be approximated by the sum of sinusoids at different frequencies. This decomposition, also referred to as the Fourier transform [Cooley and Tukey, 1965] is defined for a continuous function f as:

$$\hat{f}(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-i\omega t} dt \tag{1.1}$$

The EEG signal obtained from digital amplifiers however, is a discrete signal [Clerc, 2016]. In order to decompose it into a sum of sinusoids, the Discrete Fourier Transform (DFT) is employed [Winograd, 1978]. Considering a discrete signal f, measured in a time interval $[0,T]$ with a sampling frequency $F = \frac{1}{\tau}$ resulting in N discrete samples, the DFT of f is defined for

$k = 0$ to $k = N - 1$ as:

$$\hat{f}[k] = \sum_{n=0}^{N-1} f[n] e^{\frac{-2i\pi kn}{N}} \tag{1.2}$$

A signal sampled at a frequency F has a Fourier decomposition into sinusoids with frequencies ranging from $-\frac{F}{2}$ to $\frac{F}{2}$ with a resolution, i.e. the distance between 2 consecutive frequencies of $\frac{F}{N}$ with N being the size of the discrete signal.

Spatial Representation

In some cases, the relevant information in the EEG signal does not only lie in its temporal or spectral components. To characterize certain mental tasks, it is important to know the location of the changes in the brain. An example of this can be, the already introduced covert attention paradigm. In fact, a shift of covert attention towards the right is associated with an increase in the power of the alpha activity in the right part of the occipital lobe, and a decrease in the left part. Conversely, a shift towards the left, is associated with an increase in the power of the alpha activity in the left part of the occipital lobe, and a decrease in the right part [Tonin et al., 2013]. Knowing this, it is possible to use the spatial information to determine to which side users are shifting their attention.

Statistical Representation

Another way of describing EEG signal, is to consider it as a *Vector of Random Variables*, where each electrode is considered as a random variable of real values. In this representation, the relevant information lies in the relative behavior or dynamics of each pair of electrodes, best synthesized in the *Covariance Matrix* [Congedo et al., 2017].

For an epoch of EEG recording of N channels, the covariance matrix C is a symmetric definite positive matrix [Congedo et al., 2017] where the non-diagonal elements $(c_{ij}) = Cov(e_i, e_j)$ the covariance between random variables (EEG channels) e_i and e_j, and the diagonal elements $(c_{ii}) = Var(e_i) =$ the variance of the variable (EEG channel) e_i. In practice, the variance and covariance of the channels are empirically estimated from collected data [Kalunga et al., 2016].

As stated before, the choice of the representation for the recorded EEG signal, depends on the implemented BCI paradigm. However, they are not mutually exclusive, and in many cases it is advantageous or even necessary to consider multiple types of information and thus, to use multimodal representations of the EEG signal. In the following sections, the presented methods will be categorized according to what representation of the signal they assume.

1.5.2 Signal Preprocessing

The EEG signal obtained from measuring the brain activity is mainly composed of 2 elements: the signal containing the relevant information, and the noise [Clerc, 2016]. The relevant information designates all information related to the brain activity and the neuromarkers that correspond to the performed mental task. The noise, on the other hand, comprises all the information unrelated to the mental task. It can be generated by external sources related to the hardware or the environment (power-line, loose contact electrodes etc.), or internal sources originating from the user's body (muscle artefacts, disrupted attention etc.) [Delorme et al., 2007].

EEG signal is known to be particularly noisy [Delorme et al., 2007]. It is especially sensitive to muscle artifacts, i.e electrical activity originating from the activity of the muscles (EMG: Electromyography) and ocular activity (EOG: Electrooculography). These artefacts are particularly disruptive because their amplitude are much higher than the amplitude of the EEG.

The first step of the BCI processing pipeline is *the signal preprocessing*. This step aims at (1) removing as much noise/artefacts as possible and (2) enhancing the relevant information, in order to increase the signal-to-noise ratio. Several tools have been proposed throughout the years, in order to clean the signal.

Based on Temporal Representation

From the temporal representation of the signal, it is possible to visually detect electrodes that are not behaving correctly. For example, in situations where the contact between the electrode and skull is lost, or when the measured signal is clearly noisy [Clerc, 2016].

It is also common when using the temporal representation of EEG, to downsample the signal. This step mainly serves to reduce the dimensionality of the signal.

Based on Spectral Representation

Some sources of noise and artifacts are known to have specific frequency components (e.g the power-line frequency, 50Hz in Europe and 60Hz in USA). These kinds of artifacts can be removed from the frequency components of the EEG signal using *Frequency Filtering* [Clerc, 2016, Motamedi-Fakhr et al., 2014]. In particular:

- *High-pass* filters attenuate low frequency components such as heartbeats.
- *Low-pass* filters attenuate high frequency components of the signal, usually associated with muscle artefacts.

– *Band-pass* filters attenuate the components outside a certain frequency range, given by the cut-off frequencies.
– *Notch filters* are used to attenuate a specific frequency, like the power line frequency.

Frequency filters are also useful for separating EEG components. Some neurophysiological markers only appear in specific frequency bands (the brain rhythms). Using a band-pass filter, it is possible to keep only the relevant frequency bands for the considered paradigm (e.g. the μ and β bands for motor imagery). In the case of SSVEP, several band-bass filters can be used to only keep the EEG components at the neighbourhood of the stimulation frequencies. For ErrPs, as the ERP is known to occur in low frequency bands [Chavarriaga et al., 2014], it is possible to preprocess the signal using a low-pass filter.

Based on Spatial Representation

Preprocessing EEG signal based on its spatial representation often involves an operation known as *Spatial Filtering*. Spatial filtering consists in selecting and combining the amplitudes of electrodes, by weighting them in order to emphasize the contribution of the most relevant electrodes by removing the background activity of the brain [Lotte and Guan, 2010, Clerc, 2016]. Among the spatial filtering methods for EEG preprocessing:

– *Selection a Priori*: The simplest way to spatially filter EEG signal, consists in keeping only the electrodes placed at the regions where it is known that the brain activity is relevant. For example, when designing an SSVEP or other vision based paradigms, it is possible to only keep the occipital electrodes. In terms of spatial filter, this is equivalent to give a weight of 0 to all non-occipital electrodes.
– *Surface Laplacian*: This spatial filter aims at removing local background activity from each electrode [McFarland, 2015], by removing from each channel, the mean of the 4 neighboring electrodes with the following equation:

$$\hat{E}_i = E_i - \frac{1}{4} \sum_{j \in V_i^4} E_j \tag{1.3}$$

where V_i^4 represent the 4 neighboring electrodes, of electrode Ei
– *Common Average Reference (CAR)*: This spatial filter aims at removing from each electrode, the global background activity [McFarland et al., 1997]. It supposes that most of the common activity over the electrodes are not relevant in the considered BCI task. Thus, CAR filtering consists in removing from each electrode, the mean of all the electrodes,

with the following equation:

$$\hat{E}_i = E_i - \frac{1}{N_e}\sum_{j=0}^{N_e} E_j \tag{1.4}$$

where N_e is the number of electrodes.

Based on Statistical Representation

When the EEG is represented as a statistical model, more advanced spatial filtering methods can be employed. The goal of spatial filtering is to determine a small number of virtual channels, each one being a linear combination of the recorded channels [McFarland et al., 1997]. The difference this time, is that instead of determining these combinations a priori, statistical methods are used to automatically determine the spatial filters from the recorded data.

- *Principal Component Analysis (PCA)*: PCA is a statistical method that serves 2 main purposes. By performing an orthogonal projection (rotation of the axis) of the random variables, it allows to (1) separate the principal components of a set of random variables and (2) reduce the dimensionality (number of random variables) while keeping most of the relevant information [Lee and Choi, 2002]. Computing the new set of axes for the random variables is equivalent to diagonalize the empirical matrix of covariance of the random variables [Clerc, 2016]. The new axes (components) correspond to the eigenvectors of the covariance matrix [Clerc, 2016]. They are ordered following the amplitude of their associated eigenvalue, so that the highest eigenvalue is associated with the eigenvector (component) that contains most of the relevant information.
- *Independent Component Analysis (ICA)*: ICA is a statistical method for blind source separation [Comon and Jutten, 2010, Clerc, 2016]. It assumes that the recorded EEG signal is the result of mixing K statistically independent sources with a linear model of EEG given by the following equation for each channel $x_k, k = 1, ...K$:

$$x_k(t) = \sum_{i=1}^{K} a_{kj} s_j(t) \tag{1.5}$$

with s_j the source of the signal. The mixing coefficients of each source, a_j can be organized into a mixing matrix A so that the previous equation becomes:

$$X = A.S \tag{1.6}$$

The goal of ICA is to estimate a "demixing" matrix $W \approx A^{-1}$ in order to maximize the statistical independence between the reconstructed sources:

$$\hat{s}(t) = \sum_{j=1}^{K} w_{kj} f_j(t) \tag{1.7}$$

– *Common Spatial Pattern:* CSP is one of the most employed spatial filtering techniques for EEG [Lotte and Congedo, 2016]. In short, the goal of CSP is to determine a set of spatial filters, in a way that the variance is minimal for a class, and maximal for another one, maximizing the inter-class difference [Ramoser et al., 2000]. With the fact that the variance of the EEG signal, filtered in a certain frequency band, is equal to the power of the signal at that frequency band [Lotte and Congedo, 2016], CSP is particularly suited for BCI based on oscillatory components. For example, CSP allows to determine spatial filters that maximize the difference between SSVEP classes, when for each class, the signal is filtered around the stimulation frequency. Formally, CSP determines the spatial filters ω by maximizing and minimizing the following function:

$$F(\omega) = \frac{\omega C_1 \omega^T}{\omega C_2 \omega^T} \tag{1.8}$$

where C_i is the covariance matrix of the EEG data for class i. This problem can be solved by a *Generalized Eigen Value Decomposition*, because $F(\omega)$ is under the form of a generalized Rayleigh quotient. More details on the computation of CSP filters can be found in [Lotte and Congedo, 2016]. Finally, as much as CSP have been proven efficient for spatial filtering EEG based on oscillatory components, it completely ignores the temporal information of the signal.

– *xDAWN algorithm:* xDAWN is a spatial filtering technique, particularly designed for enhancing the SNR of Evoked Potentials [Lotte and Congedo, 2016, Rivet et al., 2009]. Formally, xDAWN determines the spatial filters by maximizing the following objective function:

$$F(\omega) = \frac{\omega S S^T \omega^T}{\omega X X^T \omega^T} \tag{1.9}$$

where S is the average time course of the evoked potential (average from recorded trials) and X is the whole EEG signal. Maximizing the previous function thus leads to maximize the power of the relevant signal and minimize the power of the whole signal plus noise.

1.5.3 Feature Extraction

Once the signal is preprocessed, the next step of the processing pipeline is to describe the signal of interest with limited and informative features. The objective of this step is to identify and select the relevant features from the whole signal, that allow to characterize the presence of neuromarkers related to the task, and best differentiate between the monitored mental states. In addition, extracting features allows to reduce the dimensionality of the data which may speed up the processing and avoid the curse of dimensionality of some classification methods [Trunk, 1979].

Based on Temporal Representation

For the BCI paradigms that are best described by their time course, instead of using all the time steps of an epoch, a simple feature extraction consists in selecting specific time points [Lopes-Dias et al., 2018]. The amplitude of the signal for each electrode at every selected time points are concatenated into a feature vector that is later being classified. This approach can for example be employed for detecting ErrPs.

Based on Spectral Representation

For BCI based on oscillatory components, a simple feature extraction technique consists in determining the power of the EEG at the frequency bands of interest, and concatenating them into a feature vector [Lotte and Congedo, 2016]. This is for example the case for SSVEP-based BCIs where a simple feature vector consists of the concatenation of the power of the signal at the frequencies of stimulation.

Based on Spatial Representation

Feature extraction based on the spatial representation usually consists in selecting subsets of electrodes that best characterize the EEG signal. Spatial filtering methods are also considered feature extraction methods [Clerc, 2016, Lotte and Congedo, 2016].

Based on Statistical Representation

Feature extraction methods based on statistical representations try to exploit relationships between electrodes. For example:

– *Canonical Correlation Analysis (CCA)*: CCA is a very effective feature extraction method for SSVEP recognition [Lin et al., 2006]. Given 2 sets (column vectors) $X = (x_1, ..., x_n)'$ and $Y = (y_1, ..., y_m)'$ of random variables, CCA consists in finding 2 vectors $a (a \in \Re^n)$ and $b (b \in \Re^m)$ that maximize the correlation $\rho = corr(a'X, b'Y)$ by solving the following problem :

$$\max_{a,b} \rho(X,Y) = \frac{a'cov(X,Y)b}{\sqrt{a'cov(X,X)a}\sqrt{b'cov(Y,Y)b}} \tag{1.10}$$

When used for SSVEP recognition, the first set of variables is the EEG signal $x(t)$ recorded from N_e electrodes with a sampling frequency of S and T sampling points:

$$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ . \\ . \\ . \\ x_{N_e}(t) \end{pmatrix}, \qquad t = \frac{1}{S}, \frac{2}{S}, ..., \frac{T}{S} \tag{1.11}$$

And, the second set is a template signal for each one of the stimulation frequencies $y_i(t), i \in [[1, ..., N_f]]$ with N_f the number of stimulation frequencies. For each frequency f_i, the template signal is built by concatenating sine and cosine waves of frequency f_i and a number of harmonics N_h corresponding to the decomposition into Fourier series :

$$y_i(t) = \begin{pmatrix} y_{i_1}(t) \\ y_{i_2}(t) \\ y_{i_3}(t) \\ y_{i_4}(t) \\ . \\ . \\ . \\ y_{i_{2N_h}}(t) \end{pmatrix} = \begin{pmatrix} \sin(2\pi f_i t) \\ \cos(2\pi f_i t) \\ \sin(4\pi f_i t) \\ \cos(4\pi f_i t) \\ . \\ . \\ . \\ \sin(2N_h \pi f_i t) \\ \cos(2N_h \pi f_i t) \end{pmatrix}, t = \frac{1}{S}, \frac{2}{S}, ..., \frac{T}{S} \tag{1.12}$$

The value of the canonical correlation (i.e after linear transformation) between X and Y_i is

called the Canonical Correlation coefficient. By concatenating the canonical correlation coefficients with each stimulation frequency, the new feature vector v_j describing an EEG epoch is given by:

$$v_j = \begin{pmatrix} CC_{j_1} \\ CC_{j_2} \\ . \\ . \\ . \\ CC_{j_{N_f}} \end{pmatrix} \tag{1.13}$$

with CC_{j_k} being the CCA coefficient between the j^{th} epoch and the k^{th} stimulation frequency.

1.5.4 Classification

After cleaning the signal from noise and transforming it into a set of descriptive features, the next step is to classify the features into a command [Lotte et al., 2018]. In general, classification consists in associating a set of features to a mental state. This step usually follows a machine learning approach requiring a training and a working phase. The training phase consists in using a large number of examples (labelled or unlabelled [Huebner et al., 2018]) to determine a function f, that for an input feature vector x, gives the associated mental state y (i.e. $f(x) = y$). This function is called "classifier". Classifiers are then applied to classify new sets of feature vectors into one of the possible mental states. Readers interested by in-depth description of the classifiers used for BCI, may refer to [Lotte et al., 2007, Lotte et al., 2018]. Here, we describe some of the algorithms, exploited in the next Chapters, that are used to classify the extracted features into mental states.

Linear Discriminant Analysis (LDA)

LDA is a classification algorithm that seeks to construct a hyperplane to separate between 2 classes [Mika et al., 1999]. Considering the training samples (feature vectors) as multidimensional points, LDA seeks to define a hyperplane maximizing the distance between the means of the 2 classes, while minimizing the interclass variance [Fukunaga, 1990]. In the case where there are more than 2 classes, the usual strategy is to determine one hyperplane between each class and

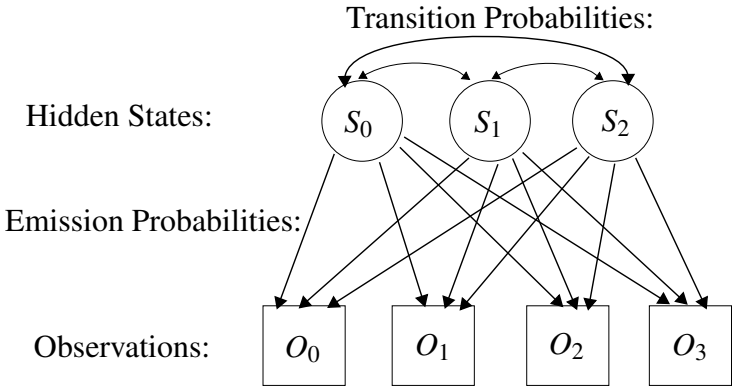


Figure 1.17: Graphical representation of a Hidden Markov Model $HMM = (N, M, A, B, \pi)$. In this example, $N = 3$ hidden states: $S_0, S_1, and S_2$ can only be characterized through $M = 4$ observation symbols: $O_0, O_1, O_2 and O_3$. The transition probabilities A represents the probability to move between each hidden state. The emission probabilities B are the probabilities of each hidden state generating one of the observation symbols.

the rest (One Versus Rest).

LDA assumes a normal distribution of the data as well as the equality of the covariance matrices for the classes [Lotte et al., 2007]. It has the advantage of having a very low computational complexity, which makes it particularly suitable for online BCI systems.

For example, using LDA to classify ErrPs consists in training the classifier to differentiate between positive (presence of ErrP) and negative (absence of ErrP) classes by building training samples using temporal features represented by the amplitude of the EEG signal at specific time points. LDA can also be used to classified SSVEP [Legeny et al., 2013]

Hidden Markov Models

A system that has a certain number of states, with probabilities of transitioning between them over time can be represented as an automaton. Under the hypothesis that the probability of being in a certain state only depends on the previous one, we define a first order Markov Model [Kemeny and Snell, 1976].

For certain classes of systems, like EEG, these sequence of states are hidden and only observed through another stochastic process of observations that only depends on the hidden states (the features). This doubly embedded stochastic process can be represented as an automaton (see Figure 1.17), called a Hidden Markov Model (HMM). Formally, a Hidden Markov Model is defined using 5 parameters:

$$\lambda = (N, M, A, B, \pi) \tag{1.14}$$

with N the number of states, M the observable symbols (vocabulary), A the matrix of state transitions probabilities, B the matrix of emission probabilities (i.e the probability of each observation being emitted by one of the hidden states), and π the initial state distribution.
HMMs allow to solve 3 types of problems [Rabiner, 1989]:

- Problem 1: Given a sequence of observations $O = O_1 O_2 ... O_T$ and a model $\lambda = (A, B, \pi)$ [14], what is the probability that the model generated the sequence: $P(O|\lambda)$?
- Problem 2: Given a sequence of observations $O = O_1 O_2 ... O_T$ and a model $\lambda = (A, B, \pi)$, what is the sequence of hidden states $Q = q_1 q_2 ... q_T$ that is the most likely to have generated or best explains O?
- Problem 3: Given observation sequence $O = O_1 O_2 ... O_T$, what are the model parameters $\lambda' = (A', B', \pi')$ that guarantee $P(O|\lambda') > P(O|\lambda)$?

For modeling brain activity using Hidden Markov Models, the mental states are associated to the hidden states. And the observed features from the EEG are associated with the observations. Using a HMM to recognize SSVEP responses is done through solving problem 1 and problem 3 from the above mentioned:

Training Phase: For each class or mental state, a Hidden Markov Model is trained by solving the Problem 3 (i.e using the available training data to determine the parameters of the HMM). In particular, training trials are turned into sequences of feature vectors corresponding to sequences of observations, and using the Expectation-Modification (EM) algorithm [Rabiner, 1989], the parameters of the HMM are estimated to best fit the obtained sequences. One of the main advantages of the HMMs, is their ability to capture temporal information from the training data, and use this information for the classification [Lotte et al., 2007].

Classification Phase: The classification phase consists in determining for a new sequence of feature vectors, the model that was most likely to have generated this sequence of features (observations). This is done by solving Problem 1. Formally, for a new sequence O, using the forward-backward procedure [Rabiner, 1989] we determine the likelihood $L_i(O)$ of the sequence being generated by the model i. The recognized class corresponds to the model that is the most likely to have generated the sequence.

14. This is a compact representation of a HMM as the number of states and observations can be determined from transition and Emission Matrices

1.6 Conclusion

In this Chapter, we have introduced the field of Brain-Computer Interfaces and described the steps that allow to translate brain activity into computer commands. We have started by describing the fundamental concepts about the brain architecture, and highlighted how the location of the brain activity depends on the nature of the mental task or the sensory stimuli. We presented the principle of EEG measurement, and presented all the steps of the BCI processing pipeline.

We have also examined the different representations of electroencephalography, specifying the types of information each one emphasizes. Then we have reviewed examples of preprocessing techniques to de-noise the recorded signal and to separate its main components. We described feature extraction methods and classification algorithms based on these different representations.

CHAPTER 2

BRAIN-COMPUTER INTERFACES AND AUGMENTED REALITY

Abstract:

This Chapter reviews the state of the art of using Brain-Computer Interfaces (BCIs) in combination with Augmented Reality (AR). First, it introduces the field of AR and its main concepts. Second, it describes the various systems designed so far combining AR and BCI categorized by their application field: medicine, robotics, home automation and brain activity visualization. Finally, it summarizes and discusses the results of the survey.

2.1 Introduction

Brain-Computer Interfaces and Augmented Reality are two fields that can be combined for interaction and/or visualization purposes. On the one hand, AR-based systems usually rely on Head Mounted Displays (HMD), that can be used in scenarios requiring hands-free interaction [Blum et al., 2012]. BCI paradigms can provide such means of input, either to interact with virtual [Faller et al., 2017] or real objects [Takano et al., 2011]. On the other hand, BCIs can take advantage of AR in order to enable interaction with the real world. AR can also provide interesting ways of displaying feedback by integrating it in the real world environment.

This Chapter aims at providing some notions about AR, before presenting previous and related work regarding the combination of BCI and AR. In particular, we highlight for every previous work, the purpose, the type of AR, the hardware display as well as the utilized BCI paradigm.

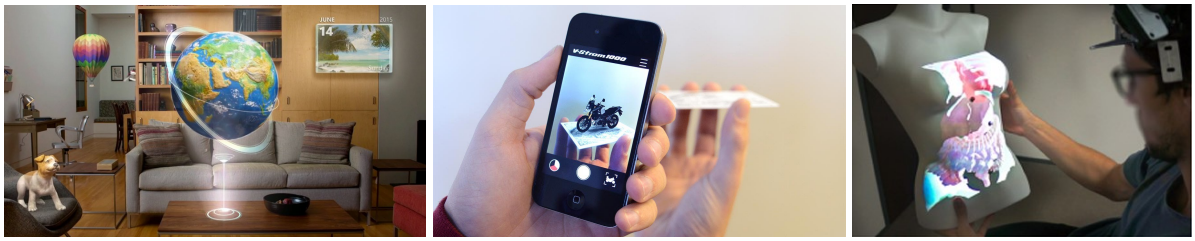


Figure 2.1: Different types of AR [1]. (Left), virtual elements directly superimposed in the user's field of view (Optical See-Through AR). (Middle), a virtual element added to a video stream recorded on a phone (Video See-Through). (Right), virtual elements superimposed on a physical object (Spatial Augmented Reality) [Cortes et al., 2018].

2.2 Notions of Augmented Reality

As stated in the Introduction, Augmented Reality is the integration of virtual objects and information in the real world in real-time [Zhou et al., 2008, Azuma, 1997], with 3 main characteristics:

1. The combination of virtual and real elements.
2. The real-time interaction.
3. The 3D registration of the virtual content in the real environment.

For this, the rendering and display of virtual elements has to adapt to the user point of view in real-time. Augmenting Reality can generally be obtained in 3 different ways [Azuma, 1997, Azuma et al., 2001] (see Figure 2.1):

1. Optical See-Through (OST) AR: in which the virtual content is directly displayed in front of the user's eyes onto a semi-transparent screen (e.g., Microsoft Hololens)
2. Video See-Through (VST) AR: in which real images are shot by the camera of a device (tablet, phone, hmd) before being visualized through a screen, augmented with virtual information;
3. Projected AR (a.k.a. Spatially Augmented Reality): in which virtual content is projected into the real environment object

An essential part of any AR system is the ability to collocate virtual and real objects, which is known as registration. Afterwards, tracking allows to properly render the change in virtual

1. Image Sources: (Left) https://www.wired.com/2015/04/microsoft-build-hololens/ (Middle) https://www.youtube.com/watch?v=J8qDRV1RxTw

objects according to the position of the camera and thus, ensuring their credible integration into the real world [Zhou et al., 2008]. Registration of virtual elements can be done using fiducial markers placed in the real environment, through pattern recognition techniques to identify real objects or with active sensors [Azuma, 1997].

One popular way of achieving the tracking, consists in using the Simultaneous Localization And Mapping (SLAM) algorithms [Billinghurst et al., 2015] related to the resolution of the problem of enabling a robot to simultaneously discover its surroundings and infer its position [Thrun and Leonard, 2008]. Originally designed for robots' navigation [Dissanayake et al., 2001], it has been adapted for use in AR [Davison et al., 2007] as it allows the tracking of objects in unknown environments [Billinghurst et al., 2015].

One of the major challenges for AR is "interaction". It is necessary to provide the user with means to act on the virtual elements [Zhou et al., 2008] and to manipulate them. However, being in the context of wearable computers, new ways of interaction, different from mouse and keyboard, have to be employed.

So far, interacting in AR has mainly been done through voice commands and hand gesture recognition [Irawati et al., 2006] (as with Microsoft's Hololens), gaze tracking [Höllerer and Feiner, 2004] or with physical buttons [Schall et al., 2009] (as with Google Glasses). BCIs could particularly contribute to AR-based systems interaction means, especially on visual selection tasks that can be done via SSVEP or P300 for example [Gergondet et al., 2011, Lenhardt and Ritter, 2010].

2.3 Combining BCI and AR: Application Fields

Combining AR and BCI could potentially be applied to most topics where BCIs can, e.g. assisting disabled people, entertainment, sports. There may be different reasons why to combine AR and BCI. First, from a BCI point of view, AR offers new ways to integrate feedback in real world environment, thus, bringing new interaction possibilities and enhancing the user experience. Second, from an AR point of view, BCIs offer new hands-free paradigms for interaction with physical and virtual objects as well as new physiological information about the user's mental state, allowing to create more adaptive scenarios.

This section presents the state of the art of combining BCI and AR systems categorizing them by application field (Medicine, Robotics, Home Automation, Brain Activity Visualization).

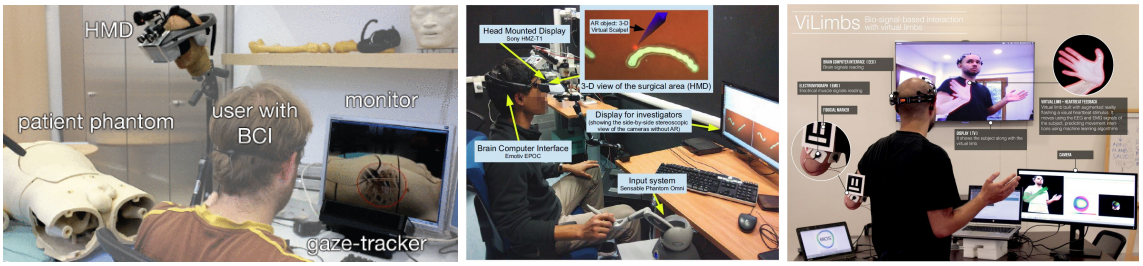


Figure 2.2: Examples of works combining BCI and AR for Medicine. (Left) represents an AR interface for helping surgeons have multiple levels of depths view, using a BCI as an interface to switch between the levels [Blum et al., 2012]. (Middle) represents a system based on BCI and AR for training surgeons keep engaged while operating. The BCI is used to measure the level of engagement [Barresi et al., 2015]. (Right) represents a system based on AR and BCI for enhancing the mirror therapy, replacing a missing limb with a virtual one, controlled with EEG and MEG [Correa-Agudelo et al., 2015].

2.3.1 Medicine

Three main types of applications combining AR and BCIs for medicine can be identified (see Figure 2.2):

1. Surgeons aid or training
2. Psychological treatments
3. Disabled people assistance

An attempt to aid surgeons during operation is the work of Blum et al. [Blum et al., 2012] who developed a Video See-Through Head Mounted Display (VST HMD) AR system granting "X-ray Superman-like vision" to surgeons in order to let them have more in-depth vision of patients under surgery (see Figure 2.2 Left). The goal of this application was to combine a BCI with a gaze-tracker, the latter selecting the area where to zoom-in and the former being used to control the level of zoom. The main utility of using a BCI in this context, is that surgeons act in a totally hands-free context, as their hands are sterilized and hence, cannot be used to interact with the AR-System [Blum et al., 2012]. However, their final setup relied on EMG instead of EEG.

When it comes to help surgeons, this can either be done by providing them with tools to use during operations, or to provide ways for them to train before to operate. This has been done by Barresi et al. who developed a prototype called BcAR [Barresi et al., 2015]. They combined BCIs and AR feedback in order to train surgeons for Human-Robot-Interaction based surgeries. In BcAR, surgeons train for robot-assisted laser microsurgery. They have to manipulate

a "retractable" scalpel represented by a haptic arm. AR feedback, displayed through a Video See-Through Head Mounted Display (VST HMD), is used to show them their attention level – measured through the BCI – represented by the length of the scalpel, so that they can adapt it (see Figure 2.2 Middle). The goal of the system is to teach surgeons keep their concentration during the whole time of the operation.

Another therapy that has been enhanced by combining AR and BCI is the “exposure therapy”. To cure patients from phobias and anxiety, Acar et al. developed an EEG based system to help patients overcome their fear [Acar et al., 2014]. The AR system consisted of a smartphone, displaying a camera view augmented with the entity the user feared (such as insects), to help them confront it. EEG was measured in order to determine the efficiency of this AR-enhanced exposure therapy.

As stated before, BCIs and AR can also be combined in order to enhance psychological therapies. [Correa-Agudelo et al., 2015] developed ViLimbs, a computer screen based AR-BCI for phantom limb pain treatment. In this system, patients are placed in front of a wide screen displaying a facing camera stream (see Figure 2.2 Right). Thanks to a fiducial marker placed on the beginning of the missing limb, patients have an image of themselves with both arms, allowing them to move the missing one from painful positions. It is hence, an enhanced version of the mirror therapy. Brain and muscle activity are used to determine user’s motion intent to allow him to move his virtual limb. Despite using EEG, authors’ prototype relied 80% on myoelectric signals and far less on Motor Imagery [Correa-Agudelo et al., 2015].

Another example of medical application combining BCIs and AR concerns assistive technologies, particularly electric wheelchair control. This has been explored by Borges et al. [Borges et al., 2016] who are designing an environment to allow disabled people to safely learn how to drive a wheelchair. Among different modalities to drive the wheelchair, they designed an SSVEP-based solution to control the direction. The goal of AR in this system was to be able to provide different driving scenarios by integrating virtual obstacles to the real world scene while still ensuring users’ safety.

2.3.2 Robotics

BCIs and AR have particularly been used in the field of Robotics for 2 main purposes (see Figure 2.3):

1. Explicitly steering or controling robot agents
2. Manipulating robotic arms

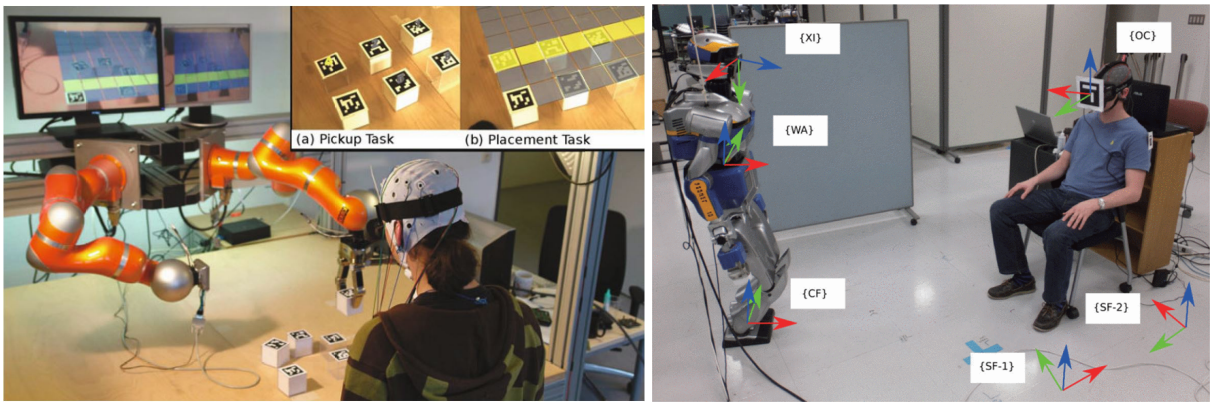


Figure 2.3: Examples of works combining BCI and AR for Robotics. (Left) represents a BCI and AR system for controlling robotic arms for a drag and drop task of cubes, using P300 [Lenhardt and Ritter, 2010]. (Right) displays a system using BCI and AR for steering a humanoid robot [Petit et al., 2014].

It is possible through AR, to provide a first-person view of the real world, augmented with contextual visual commands. This has been demonstrated by works like Escolano et al. who developed a P300-based AR system to control a mobile robot [Escolano et al., 2011]. The robot was in a different room, equipped with a camera displaying a first-person view on a computer screen, augmented with a P300 menu to control it.

A similar work had also been done by [Gergondet et al., 2011] who proposed a system to steer a robot using SSVEP. Their system allowed users to control a robot equipped with a camera displaying the augmented robot's view on a computer screen. But in this case, the menu consisted of four flickering commands. In addition to merely steer the robot, it was possible to select different speeds.

Petit et al. developed a robot navigation system to allow users to interact with a robot [Petit et al., 2014]. Thanks to a fiducial marker placed on the user's VST HMD, the user can make the robot come towards him. Then, a body part selection happens with fiducial markers placed on different parts of the user's body beginning to flicker so that they can be selected through SSVEP for the robot to interact with (see Figure 2.3 Right).

BCIs and AR have also been used to control robotic arms through goal selection (shared control) rather than step-by-step control [Tonin et al., 2010]. This has notably been done by Lenhardt and Ritter [Lenhardt and Ritter, 2010] who have used a P300 oddball paradigm in order to make a robotic arm move real objects on a table. The objects were 5 cubes tagged with AR markers that had 3D virtual numbers appearing on top of them when seen through a VST HMD.

The numbers were highlighted in a random order to elicit a P300 response when the user wanted to select one of them (see Figure 2.3 Left).

When an object was selected, a grid appeared on the table. Each case representing a possible target destination that was also selected through the P300 paradigm. After the selection of both target object and destination, the robotic arm performed the motion.

Another robotic arm control project has been achieved by Martens et al. They designed a robotic arm for two tasks [Martens et al., 2012]. The first consisted to select and move objects through P300 paradigm. The 'stones' to move were augmented when seen through a VST HMD so that the user could focus on the stimuli. The second task was to control the robotic arm to insert a key in a keyhole and was done through the augmentation of the HMD view with four SSVEP commands.

Lampe et al. have used Motor Imagery (MI) for the purpose of controlling a robotic device present in a different location than the user [Lampe et al., 2014]. The robot was equipped with two cameras, one for hand view and the other for the scene view, and both displayed on a computer screen. Whenever a selectable object entered the field of view, it was augmented so that the user could select the object to grasp through MI, and the robotic arm autonomously grabbed it. In this case, three commands were sent through Motor Imagery: left, right, to select which object to grasp, and confirmation. These commands respectively corresponding to left or right finger tapping and toe clenching.

[Zeng et al., 2017] proposed an AR-based system combining EEG and a gaze tracker to control a robotic arm. Their system involved the control of a robotic arm to which they could send commands using a gaze tracker, by selecting the location to move the arm to. The BCI was used to differentiate between rest and control states using the Motor Imagery paradigm. EEG was recorded using a 14 channels Emotiv EPOC+, and the AR content was displayed using a VST approach, displayed on a computer screen.

Recently, [Wang et al., 2018] proposed an AR BCI system to control a flying drone through a VST-HMD. In their system, SSVEP was used in combination to head movement detection for sending 4 commands. The movement of the head was used as a switch to discriminate between transition states, and control states. Their apparatus involved a 14 channels Emotiv EPOC EEG device and a HTC Vive HMD.

2.3.3 Home Automation

Another application field deals with the ability to control smart environments, whether it is to provide automated comfort mechanisms or assistive control to manipulate home appliances

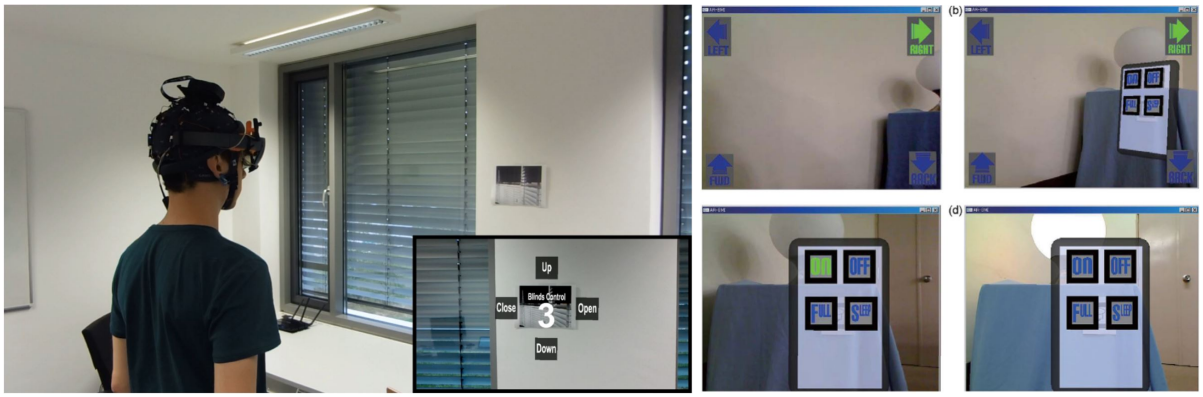


Figure 2.4: Examples of works combining BCI and AR for home automation. (Left) represents a home automation system using an OST HMD and SSVEP combined with gaze tracking for the control of home appliances [Putze et al., 2019]. (Right) displays an indirect home automation system, where the user steers a robot, with a 1st person view, and can control home appliances through P300 [Kansaku et al., 2010].

(see Figure 2.4). In this case, combining BCIs and AR is achieved through mainly two different strategies:

1. Direct Interaction.
2. Indirect interaction through a robot agent.

The first strategy has been used by Takano et al. in a P300-based AR-BCI system to control multiple devices at home [Takano et al., 2011]. They tagged house appliances with AR markers which, when seen-through an Optical See-Through (OST HMD), make a control panel appear over them. The P300 paradigm is then used to select the command to execute.

[Putze et al., 2019] proposed a home automation system enabling the control of different aspects of the house (window blinds, television, music player) using an AR-BCI system. The provided an AR interface through an OST-HMD (Microsoft HoloLens) displaying flickering targets to enable the interaction using SSVEP.

Indirect interaction has been proposed by Kansaku et al. [Kansaku et al., 2010], with a system that allows users to control a distant robot in a house environment through brain activity. The robot was equipped with a camera displaying a video stream of its environment where appliances were tagged with fiducial markers (see Figure 2.4 Right). When one of them entered the robot's field of view, a control panel was displayed, allowing users to control it.

2.3.4 Brain Activity Visualization

BCIs can be used in AR for brain activity visualization purpose (see Figure 2.5). The objective of these systems can be either:

1. Neurofeedback
2. Pedagogy

Moreover, AR can offer a natural way to display how the brain works and integrate it in real life context.

The notion of neurofeedback is an essential part of the training for BCI use [Lotte and Jeunet, 2015]. Neurofeedback has been provided in AR either by projecting it on real life objects, or displaying it directly on the representation of the user. [Mercier-Ganady et al., 2014] developed an application called MindMirror using AR for the purpose of neurofeedback. The system consisted of a smart mirror - an LCD Screen with a facing camera - displaying the user in a somehow X-Ray vision way (see Figure 2.5 Left) showing him/her the activated areas of his/her brain through EEG measurement. More precisely, the system displayed the distribution of the electrical potential over the surface of the virtual brain.

[Potts et al., 2019] proposed a new approach combining AR and BCI using neurofeedback for Meditation. Their system named *ZenG*, based on the concept of Zen Gardening, aimed at improving the users levels of relaxation and self awareness. The AR content was displayed on a OST-HMD system (*Magic Leap*). EEG was recorded using a 4 channels Muse 2014 headset, and the normalized levels of alpha power, was used as an estimate of the user's level of relaxation. Their approach was validated through a preliminary study on 12 participants.

Frey et al. developed a projected AR system called Teegi [Frey et al., 2014] which consists of a tangible figurine on the head of which, the recorded EEG of the user is displayed (see Figure 2.5 Right). The goal of Teegi was educational as it was designed for people to understand how EEG works.

2.3.5 Research Studies

Some works do not totally fall in one of these previous categories. They are proof of concepts and feasibility/research studies.

It is the case for the system of Faller et al. who developed a proof of concept of SSVEP-based BCI to control a virtual character augmented on a real table [Faller et al., 2017]. Their system included a VST HMD device, and the users' goal was to make the character move through a series of points represented by flickering checkerboards.

Figure 2.5: Examples of works combining BCI and AR for brain activity visualization. (Left) represents the *Mind Mirror* a system displaying the user's brain activity in a smart mirror for neurofeedback [Mercier-Ganady et al., 2014]. (Right) displays an example of SAR combined with BCI for brain activity visualization. The brain activity of the user is diplayed on the brain of *Teeg* a pedagogical robot for teaching EEG [Frey et al., 2014].

Another feasibility study was performed by Uno et al. who wanted to determine the effect of an uncontrolled real space background on the performance of a P300-based BCI [Uno et al., 2015]. Their preliminary results showed no effect of real space background on the selection accuracy, thus encouraging the use of combined AR-BCI applications.

Chin et al. developed a prototype in which users could reach and grasp virtual objects augmented on a real table [Chin et al., 2010]. The user's hands were augmented with virtual ones that he could control through Motor Imagery. The whole scene was displayed on a computer screen and no impact of AR was found on MI performance.

Another type of applications has made use of fNIRS in the context of wearable devices. Afergan et al. developed a fNIRS-based BCI called Phylter [Afergan et al., 2015]. Used in combination with Google Glasses, their system helped prevent the user from getting flooded by notifications. It was passively analyzing user's cognitive state to determine whether or not he/she could receive notification. The decision was based on the level of cognitive workload of the user determined after training the classifier on different user's states.

Still using fNIRS-based BCIs, [Shibata et al., 2014] presented a prototype of a Google Glass application called Zero Shutter Camera which consisted on a passive photo trigger, based on user's mental workload. The system took the predicted user mental state as input and automatically triggered a camera snapshot at 'special moments' estimated when user's mental workload was

above a threshold determined through training.

In a different context, [Angrisani et al., 2018] proposed a wearable single electrode AR-BCI system, for the monitoring of complex industrial machines. In their prototype, technicians could handlessly require information about some components of the industrial machine by sending SSVEP commands. Their apparatus involved a 2 channels dry electrode system (*OlimexEEG-SMT*) and OST AR glasses (*Epson Moverio BT-200*).

2.4 Discussion

Table 2.1 summarizes the previous works combining AR and BCIs according to the BCI paradigm, the type of AR, type of display, BCI sensor, field and objective.

This table shows first that most of the time, augmentation is done either through computer screens or HMDs, and that only a few number used Optical See-Through AR. The reason for this may be that computer screens are convenient for prototyping and the HMDs very intuitive, enabling more mobility for users. However, since screen-based AR prevents users from moving, the state of BCI development so far, also prevents them from moving with HMDs since paradigms and scenarios are designed to avoid muscle artifacts.

As combining AR and BCI is relatively new, the question of mobility did not seem to be discussed in most of the papers using HMDs. However, the development and improvement of BCI technology, notably developing filtering methods to efficiently remove muscle artifact would be a prerequisite for using BCIs as AR interaction tool to its full potential.

Regarding BCI paradigms, although a number have been considered, SSVEP and P300 paradigms are the most used ones. This popularity could be due to the graphical aspect of the augmentation, as AR is based on displaying graphical virtual elements on the users' field of view, hence, vision-based paradigms are well suited for selection tasks.

However, it is important to explore more deeply the effect of AR on BCI performances, not only from the system's performance point of view but also in terms of user interfaces.

It is noticeable from Table 1 that most of the works relied on active BCI paradigms (including reactive). They were mostly used for manipulation and voluntary control of physical or virtual objects. Passive BCIs could also be more deeply studied in the context of AR.

Finally, it seems necessary to consider AR-BCI systems from a Human-Computer Interaction perspective to evaluate and improve them. In addition, more and other fields of application could study and benefit from combining AR and BCIs in the future.

Table 2.1: Overview of previous systems combining AR and BCIs. CS: Computer Screen; VST: Video See-Through; HMD: Head Mounted Display; OST: Optical See-Through; HA: Home Automation; PoC: Proof of Concept; M: Medicine; BAV: Brain Activity Visualization; SAR: Spatially Augmented Reality. N.A: Proof of concept, no AR implemented. M.W:Mental Workload.

Work	BCI paradigm	AR type	AR display	BCI sensor	Field	Objective
[Escolano et al., 2011]	P300	VST	CS	EEG	Robotics	Robot steering
[Lenhardt and Ritter, 2010]	P300	VST	HMD	EEG	Robotics	Robot arm control
[Takano et al., 2011]	P300	OST	HMD	EEG	HA	Direct HA
[Kansaku et al., 2010]	P300	VST	CS	EEG	HA	Indirect HA
[Uno et al., 2015]	P300	VST	CS	EEG	PoC	Feasibility Study
[Martens et al., 2012]	P300/SSVEP	VST	HMD	EEG	Robotics	Robot arm control
[Borges et al., 2016]	SSVEP	N.A	N.A	EEG	M	Wheelchair control
[Gergondet et al., 2011]	SSVEP	VST	CS	EEG	Robotics	Robot steering
[Petit et al., 2014]	SSVEP	VST	HMD	EEG	Robotics	Robot steering
[Faller et al., 2017]	SSVEP	VST	HMD	EEG	PoC	Virtual char. control
[Wang et al., 2018]	SSVEP	VST	HMD	EEG	Robotics	Control of a drone
[Putze et al., 2019]	SSVEP	OST	HMD	EEG	HA	Direct HA
[Angrisani et al., 2018]	SSVEP	OST	Glasses	EEG	PoC	Industrial inspections
[Lampe et al., 2014]	MI	VST	CS	EEG	Robotics	Robot arm control
[Chin et al., 2010]	MI	VST	CS	EEG	PoC	Virtual hand grasping
[Correa-Agudelo et al., 2015]	MI/EMG	VST	CS	EEG	M	Phantom pain therapy
[Blum et al., 2012]	EMG	VST	CS	EEG	M	Surgeons assistance
[Barresi et al., 2015]	MI/EMG	VST	CS	EEG	M	Surgeons training
[Zeng et al., 2017]	MI/Gaze Tracking	VST	CS	EEG	Robotics	Robot arm control
[Acar et al., 2014]	Raw Data	VST	Smartphone	EEG	M	Phobia therapy
[Mercier-Ganady et al., 2014]	Raw Data	VST	CS	EEG	BAV	Neurofeedback
[Frey et al., 2014]	Raw Data	SAR	Puppet	EEG	BAV	Education
[Potts et al., 2019]	Alpha Level	OST	HMD	EEG	Neurofeedback	Meditation
[Afergan et al., 2015]	MW	OST	HMD	fNIRS	PoC	Proof of Concept
[Shibata et al., 2014]	MW	OST	HMD	fNIRS	PoC	Proof of Concept

2.5 Conclusion

This Chapter presented the state of the art of combining Brain-Computer Interaction and Augmented Reality. It first introduced the field of AR which can be divided into Optical See-Through, Video See-Through and Projected AR. Then it presented the previous works combining AR and BCIs in the fields of medicine, robotics, home-automation, brain activity visualization as well as proofs of concept or research studies.

This survey showed that most of the previous works made use of P300 or SSVEP paradigms in VST setups, that EEG was the most employed brain sensing technology and that robotics was the field with the highest number of applications. Combining AR and BCIs seems useful in scenarios favoring hands-free interaction.

Considering these information, in this work we study the possibility to use different BCI paradigms in AR namely: SSVEP and ErrPs. The objective being to assess the possibility to detect them in the EEG signals of users in AR as well as evaluate the effect of movements on their detection accuracy. We also consider the promising application case of combining BCI and AR: Home Automation.

In the next Part, before going further on the combination of BCIs and AR, we begin by proposing a new definition of BCIs from a Human-Computer Interaction (HCI) perspective.

PART II

On the Definition of Brain-Computer Interfaces

CHAPTER 3

PROPOSING A NEW DEFINITION OF BCI FROM A HUMAN-COMPUTER INTERACTION PERSPECTIVE

Abstract:

This Chapter aims at proposing a new definition of BCIs from a Human-Computer Interaction perspective. As BCIs are interfaces, their definition should not include the finality and objective of the system they are used to interact with. First it discusses previous definitions of BCIs in the literature. Then it compares BCIs with another widely used Human-Computer Interface (the mouse), and draw analogies in their conception and purpose.

3.1 Introduction

In 1973 Jacques Vidal introduced the term Brain-Computer Interface [Vidal, 1973]. The "BCI" project as it was imagined back then, was meant to exploit the electrical activity arising from the brain to control a computer program. Even though other terms can be found in the literature to designate BCIs: Brain-Machine Interfaces (BMI), Direct Brain Interfaces (DBI), Direct Neural Interfaces (DNI) or Brain-Interfaces (BI). It is admitted that they all designate the same thing [Mason and Birch, 2003]: an interface between a brain and a computer (in the broad sense of the term). This is even noticeable from the fact that most of them share the "I" of "Interface", a term commonly used in Human-Computer Interaction, the field that aims to design, evaluate and implement interactive systems [Hewett et al., 1992].

The most widely recognized definition of a BCI was proposed by Wolpaw et al. in 2002: "*A Brain-Computer Interface is a communication system in which messages or commands that an individual sends to the external world do not pass through the brain's normal output pathways of peripheral nerves and muscles*" [Wolpaw et al., 2002]. The main goal of a BCI would be then, to allow users who may be suffering from "locked in" syndrome or "paralysis" to

communicate and to express their wishes to caregivers, or to operate word processing programs and neuroprostheses. The very core idea of a BCI would thus be to provide the brain with “a new, non-muscular communication and control channel for conveying messages and commands to the external world”. This brings up a core element of a BCI which is the “non-muscular” nature. A BCI derives its input solely from the brain activity.

A few years later, Blankertz stated that a BCI is "*a new augmented communication system that translates human intentions reflected by suitable brain signals, into a control signal for an output device such as a computer application or a neuroprosthesis*" [Blankertz et al., 2006]. An interesting feature of this definition is the concept of translation. In order to interact using a BCI, the brain activity has to be translated into commands or messages.

According to Daly and Wolpaw [Daly and Wolpaw, 2008], a BCI system "*enables a new real-time interaction between the user and the outside world*" specifying that the signals extracted from brain activity "*are translated into an output*" from which the user receives feedback that affects his brain activity. This definition highlights the real-time component of a BCI system and introduces the notion of closed-loop between the user and the system through the feedback.

Aggregating the previously mentioned features, Graimann et al. [Graimann et al., 2009] stated 4 criteria under which a system can be called a BCI:

1. A BCI system has to acquire its inputs directly from the brain activity.
2. The signal has to be processed and translated in real-time.
3. The user must obtain feedback from his activity.
4. The user has to send intentional commands to the system.

The last criterion though, has been qualified by a newer classification of BCIs proposed by [Zander and Kothe, 2011] who distinguish between active, reactive and passive BCIs depending on the endogenous and intentional nature of the interaction. In this context, passive BCIs should also be considered as BCIs. All these definitions across time have helped to better understand and better design Brain-Computer Interfaces. They are remarkably complementary and perfectly describe what most of the BCI systems are able to achieve today, in terms of communication and control. However, most of these definitions go beyond what can etymologically fall into the definition of an interface. The term Brain-Computer Interface should only be applied to describe the intermediary hardware and software components between the brain and the interactive system.

3.2 A New Definition of Brain-Computer Interfaces

In the Oxford dictionary [1], the word “interface” defines "*A point where two systems, subjects, organizations, etc. meet and interact*". In the context of computing, it defines “*A device or program enabling a user to communicate with a computer*”. From both of these definitions, it clearly appears that the notion of interface only makes sense when considering the items it helps to bridge. In HCI, these two items are the “user” and the “interactive system”.

A system is said to be interactive if it depends on unpredictable inputs incoming from an external environment that it does not control [Goldin et al., 2006]. In HCI, the unpredictable inputs are the user commands.

In this context, an interface is the set of hardware and software means by which the user communicates with the interactive system. The interface comprises the input device (the hardware), the algorithms and methods to process its outputs and the presentation mechanisms to display the feedback. The final objective of the system though, does not belong in the boundaries of the interface. A computer mouse for example, is made of hardware parts, comprising the plastic box, the motion sensor and the microcontroller, for estimating the user movements. It also requires a transfer function for translating these movements into motion of a pointer displayed on screen and the presentation of graphical elements (i.e. buttons) so that the user can designate elements to interact with.

According to this, a BCI can be considered as any user-computer interface, since it comprises the same set of components. Hence, arising from a HCI perspective, we propose to define a Brain-Computer Interface as follows:

Definition

A BCI is any artificial system that directly converts brain activity into input of a computer process.

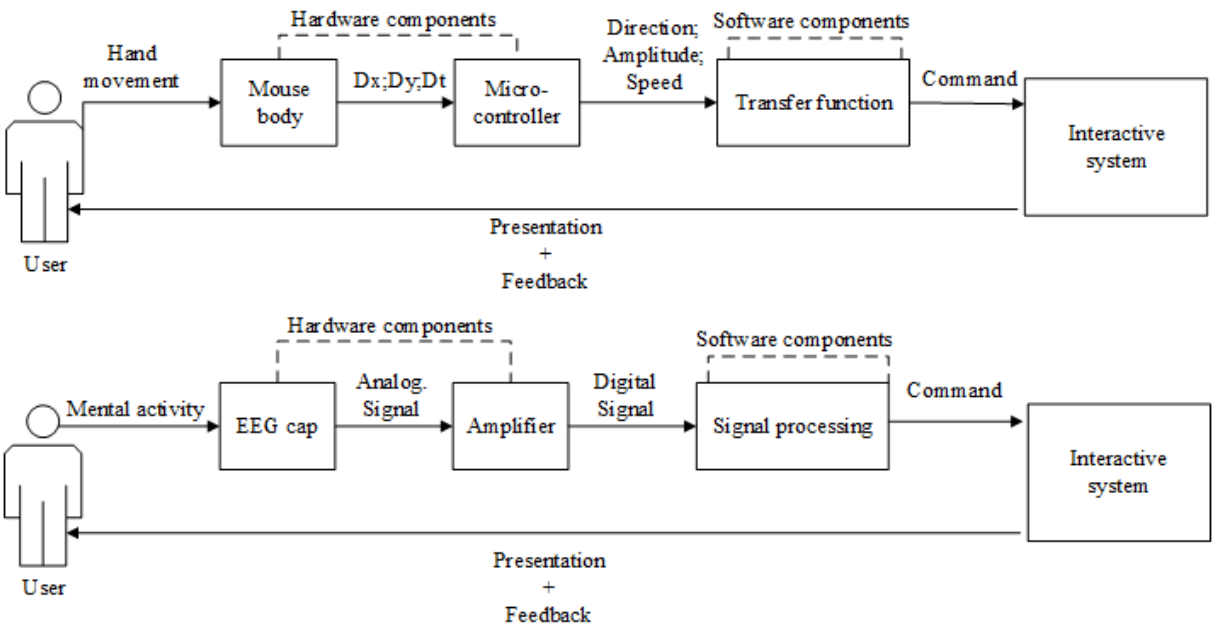


Figure 3.1: Illustration of the components comprised in a user interface. (Top) A computer mouse is made of a hardware body, a microcontroller and a transfer function. (Bottom) A Brain-Computer Interface comprises an acquisition equipment, a signal amplifier and all the methods used to transform the signal into a computer input.

3.3 Comparative Study: A Computer Mouse v.s. a SSVEP-Based BCI

In this section, in order to explain our point of view, we choose to illustrate the similarities between a Steady-State Visual Evoked Potential BCI, based on EEG, and a computer mouse. As a reminder, SSVEP is a reactive paradigm based on the property that the brain, when the subject is focused on a periodic visual stimuli of frequency f, reacts with an increase in its activity at the same frequency f. In other words, it means it is possible to determine the frequency of periodic stimulation the user is focused on. Interested readers may refer to [Vialatte et al., 2010] for more details. The summary of this analogy is illustrated in Figure 3.1 and it appears that 3 aspects should be considered.

Hardware Components: The first thing that comes to the mind of the users when they hear "computer mouse", is the plastic-made body. Similarly, when evoking a BCI is the EEG cap

1. Interface In: Oxford University Press. [online] Available at: https://en.oxforddictionaries.com/definition/interface [Accessed 20 Feb. 2019].

itself with its electrodes. In fact, both of them correspond to the entry point of the interface for the user. The second component of the mouse is the digital camera, microcontroller and the LED lens embedded in the body. It is this part that is responsible of acquiring the user's hand movements. In the case of an EEG BCI, it is the amplifier that provides one time-dependent signal per measured electrode.

Software components: In both cases of the computer mouse and the BCI, the software components hold an important part. In the mouse, the direction, amplitude and speed of the hand movement are translated into new coordinates for the cursor on the screen. In a BCI, the amplified EEG signal is processed and classified into a mental state. Depending on the mental state, and depending of the position of the cursor when clicking, a particular command is determined from the interface and sent to the interactive system.

Presentation: In both cases of the mouse and an SSVEP BCI, a graphical presentation is primordial. For the SSVEP BCI, it is essential to have external stimuli under the form of flickering targets, in order to infer different commands from the BCI. This is what makes SSVEP a reactive paradigm. For the mouse on the other hand, it is necessary to present buttons or specific areas on the screen so that the user knows where to click. In the same way as clicking on different buttons generate different commands, focusing on targets flickering at different frequencies will generate different commands.

This analogy between a mouse and a SSVEP BCI does apply to other reactive paradigms. Yet, it is possible to draw the same kind of comparison for active or passive BCIs. In the case of Motor Imagery for instance, we can easily make an analogy with a joystick or a game controller. When the subject imagines left or right-hand movement, the appropriate command is sent to the interactive system, without the need for a graphical presentation. Even though a feedback mechanism can help.

For passive BCIs, it is possible to draw a comparison with proactive and transparent interfaces. A good example of this would be a smart-home interface, where the user's behaviours and locations are transparently monitored to infer commands to send to the smart-home. Adapting a system to the user's mental state can somehow be seen as automatically turning on the light when the user moves into the room.

These comparisons and analogies are not intended to be exhaustive. They serve to illustrate that a Brain-Computer Interface, has also fundamentally to be seen as a Human-Computer Interface.

3.4 Discussion

Research and innovation in the field of BCIs is very transversal. Neuroscientists, computer-scientists, medical doctors and engineers. have different perspectives on what Brain-Computer Interfaces are.

The definition we proposed, present the advantage of being very inclusive because it does not comprise the finality of the interactive system. A similar reflection has been conducted by Jeunet et al. [Jeunet et al., 2018] when comparing Neurofeedback and Motor Imagery. They highlighted that both Neurofeedback (NF) and MI-BCI users have to learn how to regulate their neurophysiological activity, sometimes with similar features, through given feedback with different final objective. While MI-BCI consists in producing a specific EEG pattern to send a command, the goal of NF is to learn how to generate the specific pattern.

If one wants to designate the whole interaction between a user and an interactive system using a BCI to achieve a particular role, one should use the term Brain-Computer Interaction.

With our new definition of Brain-Computer Interfaces (BCI), any artificial system that involves the exploitation of brain signal, can be considered as using a BCI, instead of being the BCI itself.

3.5 Conclusion

In this Chapter, our goal was to discuss the different definitions of Brain-Computer Interfaces. All these definitions are very complementary, and reflect the transversality of the field, illustrating what it is possible to achieve with BCIs today. In an attempt to aggregate these definitions, in order to be as inclusive of new and incoming types of BCIs as possible, we proposed a new definition motivated by the etymology of the term, and with our perspective from the Human-Computer Interaction domain.

Our point of view has been illustrated through an analogy between a BCI and a more widely used interface. We believe that this new definition could enrich the discussion about what BCIs are, and that it can constitute a first step towards the promotion of Brain-Computer Interfaces as interaction media.

PART III

Investigating BCI paradigms in AR

CHAPTER 4

SSVEP: STUDY AND EVALUATION OF SSVEP IN AR

Abstract:

This Chapter aims at studying the use of SSVEP in AR. It presents two user studies conducted regarding the use of SSVEP in AR. The first experiment studies the hardware compatibility between a SSVEP-based BCI and an Optical See-Through HMD. The second experiment studies the effect of head movements on the classification accuracy of SSVEP. Each experiment is described in terms or experimental protocol, methods and results.

4.1 Introduction

In this Chapter, we study the use of a the SSVEP paradigm for interaction in an AR context. In particular, we focus on the use of an OST-HMD, knowing that our methodology can later be extended for other constitutive elements.

As mentioned before, the SSVEP is a specific brain pattern that occurs when the human visual system is stimulated by a periodic flickering stimulation. The brain responds with an activity at the very same frequency in the visual cortical area [Herrmann, 2001]. When facing multiple targets flickering at different frequencies, it becomes possible to determine which target the user is focusing on by analyzing his/her electroencephalography (EEG) signals. By associating each target with a particular command, it becomes possible to create a BCI with multiple mental commands [Evain et al., 2017]. However, such brain activities are known to be sensitive to both external and internal noise, meaning that the BCI performance can be affected for instance by electromagnetic interference with other equipment (e.g. battery) and/or by the muscular activity of the user (e.g. head movements).

The objective of this Chapter is therefore to study the feasibility of combining SSVEP and AR. Our intention is to measure the influence of both external and internal noise on the performance of an SSVEP-based BCI in AR. We conducted two user studies, assessing the

potential drops in BCI performance due to:

1. The wearing of both an EEG headset and an AR headset.
2. The movements of the user's head when observing the AR scene.

As seen before (in Chapter2), the number of works exploring the use of brain-computer interfaces in augmented reality applications remains relatively small. Although preliminary studies suggest it is possible to combine BCI with AR [Faller et al., 2017, Faller et al., 2010], studies systematically investigating the ability to combine AR and BCI remain rather rare.

The main conclusions from the study of the literature are that most of the developed systems combining AR and BCI rely on Video See-Through (VST), e.g. [Gergondet et al., 2011, Lenhardt and Ritter, 2010]. Very few researchers have explored the use of BCI with Optical See-Through AR (OST) [Takano et al., 2011] and none went beyond assessing the mere hardware compatibility between the devices.

In the following section, we address the question of the different possible sources of disruption when using SSVEP in AR with an OST HMD.

4.2 Hardware Compatibility

BCIs are known to be very noise-sensitive. Combining BCI with AR might introduce two different sources of disruption. Firstly, at the hardware level, both technologies can require head mounted devices and it is necessary to make sure that they do not interfere (*external noise*). Secondly, recording brain activity in the context of AR where users are generally free to move may also be difficult as muscle activity can provoke artifacts in the BCI recordings (*internal noise*).

Therefore, in order to assess the compatibility of SSVEP with an OST-HMD AR system, we conducted two user studies [1]. The first one, on "external noise", aimed at evaluating the effect of associating the two headsets (BCI and AR headsets) on the performance. The second study, on "internal noise", aimed at evaluating the effect of head movement on the BCI performance. In following we focus on the use of SSVEP in AR.

Before presenting the details of the experiments, it is interesting to point out that, in the literature, SSVEP responses have been reported to be successfully classified with an accuracy around 80% on screen displayed targets [Friman et al., 2007, Chen et al., 2014] and led-based

1. The experiments presented here were all conducted in accordance with the relevant guidelines for ethical research according to the Declaration of Helsinki. All the participants were briefed about the nature of the experiment and signed an informed consent form at the beginning of each experiment.

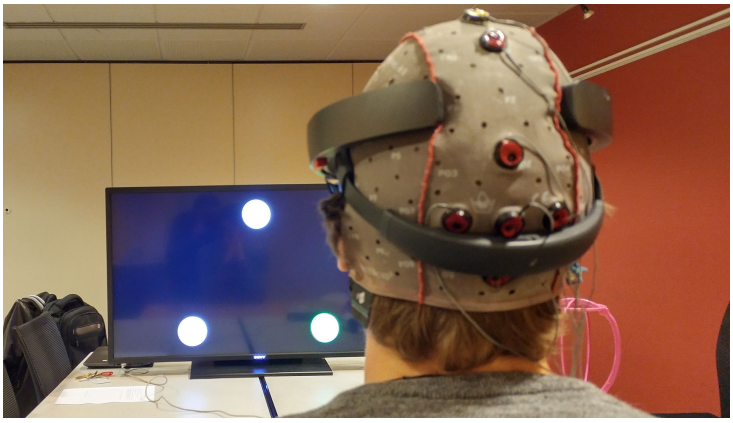


Figure 4.1: Experimental setup of Study 1: the user wears an EEG cap as well as the HoloLens. In this particular condition, the SSVEP targets were displayed on the screen.

stimulations [Hwang et al., 2012] in static contexts. Few papers addressed the issue of SSVEP selection in motion [Lin et al., 2014]. Previous studies reported a deterioration in the SSVEP classification accuracy proportional to the walking speed, but they did not consider the movement of the head which is potentially more disruptive for SSVEP due to the fact that the neck is closer to the occipital lobe (where EEG sensors are typically placed for SSVEP).

4.2.1 User Study: Influence of AR Device on BCI performance

The objective of this first study was to answer the following questions:

1. Is EEG-signal quality impaired due to the wearing of an OST-HMD?
2. Does the electrical activity produced by the display corrupt the signals?
3. Is the SSVEP paradigm effective in such a stereoscopic AR context?

These questions are assessed comparing the BCI performance while interacting with an SSVEP interface in different hardware configurations. The experimental protocol and the results of this study are presented and discussed in the following paragraphs.

Apparatus and participants

Visual stimuli (SSVEP targets) were displayed either on a 47-inch screen or on a Microsoft HoloLens (see Figure 4.1). Both display devices had the same framerate of 60 Hz. In order to keep a similar configuration between experimental configurations, participants were sat at

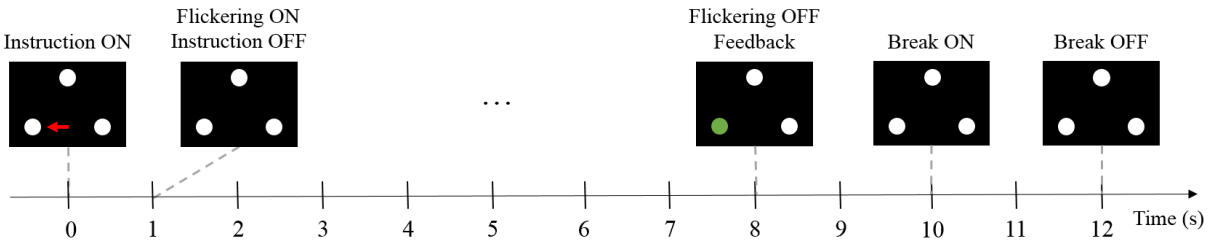


Figure 4.2: Trial timing. (t=0s) The user is instructed to focus on a particular target through an arrow. (t=1s) The arrow disappears and all targets flicker for 7s. (t=8s) The flickering stops and a sham feedback is provided (the designated target is highlighted in green in 80% of the trials). (t=10s) A two-second break starts before the next trial begins.

2 meters from the screen due to the focal distance of the HoloLens. Furthermore, the screen covered the same area of the participants' central field of view as the HoloLens.

Electroencephalography (EEG) data were recorded from a g.USBamp amplifier (g.tec, Graz, Austria) using 6 scalp electrodes (CPz, POz, O1, O2, Oz, Iz - 10-20 system) referenced to the ear and grounded to Cz. Such electrode positions enable the coverage of the visual cortex, from where SSVEP are triggered [Wang et al., 2008].

Thirteen naive participants (aged 25.9 $\pm$ 3.1 year-old, 2 women) took part in the experiment.

Experimental Protocol

Participants were equipped with the EEG cap and asked to complete 4 runs of SSVEP training. Each run lasted around 7 minutes (30 trials) and consisted in the display of 3 white targets either on the screen or through the HoloLens. The choice of the stimulation frequencies was made according to the literature, reporting which frequencies elicit the highest SSVEP responses [Pastor et al., 2003]. Each of these targets was flickering at a different frequency (10, 12 and 15 Hz). The targets were 10 cm wide circles arranged in an equilateral triangle of 46 cm side.

Participants were instructed to focus on one specific target through the appearance, for 1s, of an arrow pointing at it. After 7s, the targets stopped flickering and participants were provided with a feedback: one of the targets was highlighted in green for 2s. Given the fact that we did not perform any online classification, the feedback was sham. The correct target was highlighted 80% of the time while one of the other two was highlighted the rest of the time in order to keep the participants motivated. Then, participants had a 2s break before the next trial started. The structure of the trials is illustrated in Figure 4.2.

In total, participants had to focus on each target 10 times per run, in a random order. The

session was divided into 4 runs, with one run per condition: (C1) targets displayed on screen, no HoloLens, (C2) targets displayed on screen, HoloLens switched-off, (C3) targets displayed on screen, HoloLens switched-on, (C4) targets displayed through the HoloLens. The order of the runs was randomized across participants to avoid any potential order effect. In total, the experiment consisted in 120 trials for a duration of 22 minutes.

The classification accuracy for each condition provides insights about the potential impact of the OST-HMD on the EEG data quality and on the viability of using SSVEP in a stereocopic AR context. Our hypothesis was that the classification accuracy would not be significantly different depending on the condition, i.e., that participants would perform the same in the four conditions.

Offline EEG Data Analysis

Data were sampled at 512 Hz and processed using the OpenViBE software [Renard et al., 2010]. As stated before, the targets were flickering at 10, 12 and 15 Hz. There were 10 trials per class (each flickering frequency corresponds to a class) per run, i.e., 30 trials per run in total. To determine the performance obtained at each run, we performed an offline 10-fold cross-validation procedure. Thus, for each class of each run, 9 trials were used to train the multiclass Common Spatial Pattern (CSP [Blankertz et al., 2007]) and the Linear Discriminant Analysis (LDA [Müller et al., 2008]) algorithms while the last trial was used to test these algorithms.

This operation was computed 10 times for each class of each run. The multiclass CSP consisted in filtering the data in [9.75;10.25] Hz for the 10 Hz class, [11.75;12.25] Hz for the 12 Hz class and [14.75;15.25] Hz for the 15 Hz class and then in finding 6 spatial filters whose resulting EEG power was maximally different for one class vs. the others. The spatially filtered EEG signal power (computed on a 500ms time window, with 100ms overlap) was used to train a multiclass LDA. This LDA was then used to discriminate when the user was focusing on the targets flickering at 10, 12 or 15 Hz.

Results and Discussion

One of the samples did not follow a normal distribution. Therefore, we performed a non-parametric test for paired-samples, i.e., a Friedman test, to test this hypothesis. Given the relatively high number of samples (4 runs and 13 participants), this test could be approximated by the χ^2 test with 3 degrees of freedom (4 runs -1). In line with our hypothesis, results showed no significant difference between the runs [$\chi^2(3)=5.025$; $p=0.170$]. Results are summarized in Figure 4.3.

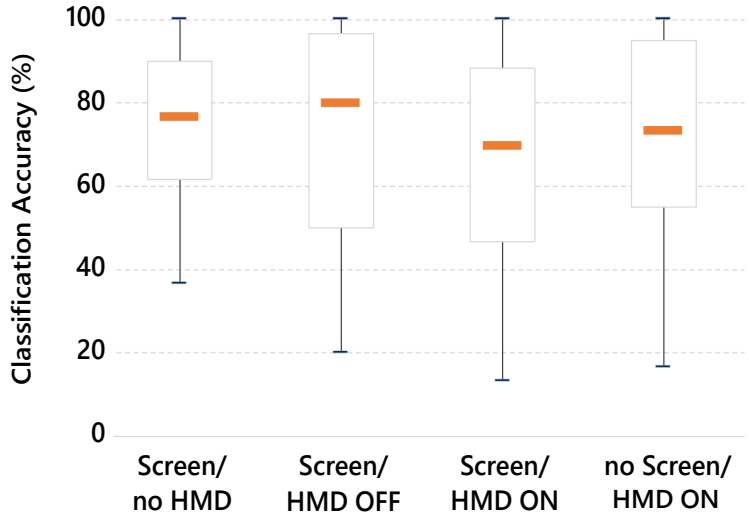


Figure 4.3: Results of Study 1: Boxplots representing the classification accuracy in % (N=13) as a function of the visual condition.

The classification accuracy in all of the conditions' was between 75% and 80% which is consistent with the literature [Friman et al., 2007, Chen et al., 2014]. It also confirms the previous results of [Takano et al., 2011] on the possibility to combine AR and BCI headsets.

Concerning user experience, participants reported that the HoloLens felt heavy after a while. Also, they reported some reflections when it was switched-on. These reflections appeared to be uncomfortable for some of them when focusing on the targets on the screen.

4.2.2 User Study: influence of Users' Movements on BCI Performance in AR

One of the main advantages when using an OST-HMD, is the possibility for the user to move while interacting with the system. Particularly, the user generally has to move his/her head in order to interact with the AR environment.

Thus, the objective of this second study was to evaluate the impact of head movements (source of internal noise) on the BCI selection performance in order to answer the questions:

1. Is it possible to perform a SSVEP selection on moving targets in AR?
2. Is head movement influencing the BCI performance?

For this matter, we use the notion of movement intensity designating a combination of amplitude and speed. A low-intensity movement is performed at low speed and low amplitude while a high-intensity movement is fast and wide. In order to assess the possibility of using SSVEP-based BCIs on moving targets, we compared user performance on selecting targets over different movement intensities.

The following paragraphs detail the implemented experimental protocol as well as the results of the experiment.

Apparatus and Participants

Visual stimuli were displayed again on a Microsoft HoloLens. The EEG apparatus was the same as in the previously described experiment. In order to ensure consistency over the results of the participants, they were asked to stand up on the same spot and only to move their head to follow the moving targets. In this experiment, we wanted to evaluate a worst real case scenario where the users would not be able to anticipate the movements of the targets, thus participants were not informed of the direction the targets would follow.

Fifteen participants (aged 22.8 $\pm$2.6 year-old, 1 woman) took part in the experiment.

Experiment Protocol

After the installation, participants were asked to perform 4 runs of SSVEP training. Each run lasted around 7 minutes and consisted of 36 trials. In this case, a trial consisted in the display of 3 moving white targets through the HoloLens. Each of these targets was flickering at a different frequency (10, 12 and 15 Hz). They were represented as 10 cm wide circles arranged in an equilateral triangle of 46 cm side in order to keep the same configuration as the previous study. Similarly, the structure of a trial and the feedback strategy were also identical to the one described in Figure 4.2.

During each run, participants had to focus on each target 12 times under a different configuration of movements, submitted in a random order. There were 4 conditions of head movements: (C1) A test condition with no movement (static targets), (C2) a low-intensity movement configuration with a maximum amplitude from left to right of 40°, (C3) a medium-intensity movement configuration with a maximum amplitude of 100° and (C4) a high-intensity movement with a maximum amplitude of 160°. In this case, the movement amplitude represents the rotation angle of the head.

The chosen movement form (path of the targets) was a lemniscate of Bernoulli (see Figure

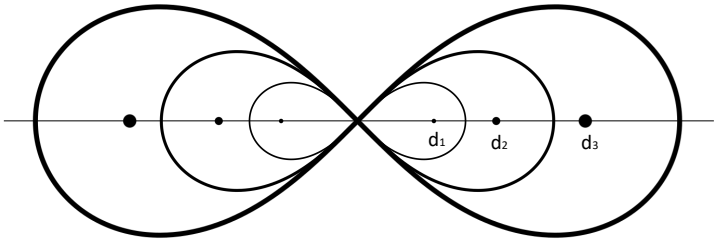


Figure 4.4: Different Movement intensities and targets motion used in user study 2: Lemniscates of Bernoulli with half focal distances $d_1(40°) < d_2(100°) < d_3(160°)$ The path of the targets lays in a sphere centered on the user.

4.4). In addition to being symmetric, this path contains different horizontal and vertical orientations of movement. The lemniscate of Bernoulli is smooth and prevents from sudden directions changes, reducing potential artifacts arising from discontinuity. The direction to perform the path was counterbalanced across each class. As the user had to focus on each class 12 times, he/she had to perform the path 3 times from each of the 4 possible initial movements i.e. beginning from (1) top right, (2) top left, (3) down right and (4) down left. This counterbalancing was done to avoid biases that might have appeared due to the repetitive execution of the same head movement which could have biased the classification.

Finally, using the parametric representation (Equation 4.1), it was possible to implement different movement amplitudes by changing the half focal distance of the lemniscate d_i (see Figure 4.4) while ensuring to perform the whole movement in 7 seconds, implementing the notion of movement intensity. In other words, a small lemniscate was performed in 7s (at low speed) and a larger one also in 7s (at a higher speed).

$$\begin{cases} x(t) = d\sqrt{2}\,\dfrac{\sin\psi(t)}{1+\cos^2\psi(t)} \\ \\ y(t) = d\sqrt{2}\,\dfrac{\sin\psi(t)\cos\psi(t)}{1+\cos^2\psi(t)} \end{cases} \quad \text{with} \quad \psi(t) = 2\pi\frac{t}{7} - \pi \tag{4.1}$$

Results and Discussion

According to the results of the literature [Lin et al., 2014, Fatourechi et al., 2007], our hypothesis was that head movement would deteriorate the classification accuracy of SSVEP responses. Particularly as this paradigm is exploited from recordings on the visual cortex at the back of the head, and as head movements solicit the neck, the strongest muscle artifacts should appear in that area, close to the EEG electrodes. Accordingly, the results of the experiment show a clear degradation of the BCI performance together with the increase in movements intensity (see Figure 4.5). Thus, average performance drops from 78%, when participants do not move, to 41% (which is around chance level [Müller-Putz et al., 2008a]) for the maximal movement amplitude.

Moreover, performing a statistical test on the data showed that this difference was statistically significant at 95% of confidence (Friedman test was performed as in section 8.6). However, after performing Post-Hoc test through multi-comparisons between conditions after Friedman test, it appeared that the difference between the “No Movement" and the “Low Intensity Movement" conditions was not significant (see Figure 4.5).

In addition to the statistical analysis, a study of the forms filled by the participants right after the experiment, revealed that a large majority of participants had difficulties following the targets in the “High Intensity Movement" condition considering it an extreme condition. Thus, the deterioration of performance could also be partially due to the inability to accurately track the targets.

In a nutshell, the results of this experiment suggest that it is possible to use SSVEP as a selection technique in the context of AR with small head movements (around 40° of amplitude). They also show a significant deterioration of performance with higher intensity movement (100°) but still above chance level [Müller-Putz et al., 2008a], whereas extremely high intensity movement prohibits the use of SSVEP.

4.3 Global Discussion

Combining Brain-Computer Interfaces with Augmented Reality raises a lot of scientific and technological challenges. It poses theoretical questions regarding the design of effective interaction paradigms but it also involves feasibility studies regarding notably the compatibility of these two complex technologies. In addition, the design of AR-BCI systems is also strongly related to the chosen BCI paradigm and the designed UI. In the case of SSVEP it involves the

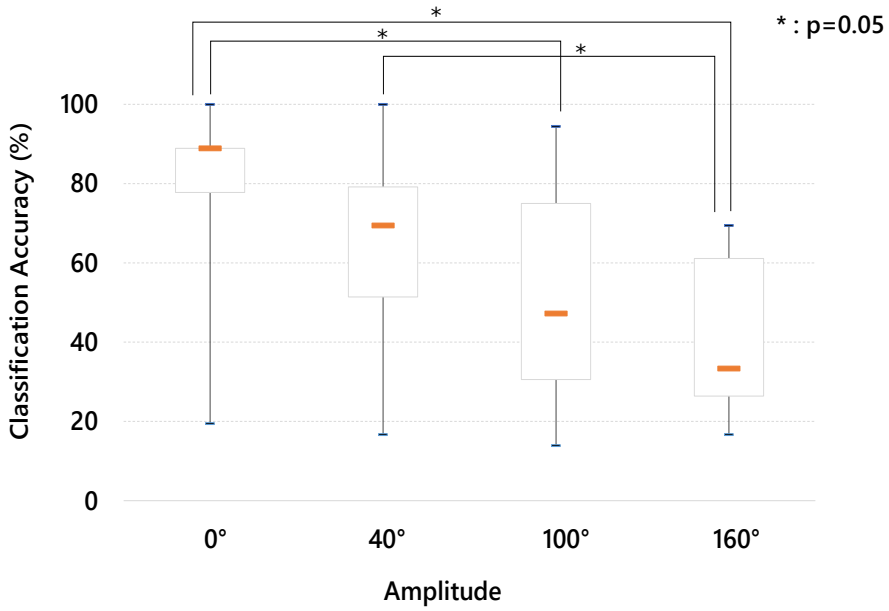


Figure 4.5: Results of Study 2: Boxplots representing the classification accuracy in % (N=15) as a function of the movement intensity.

placement of flickering targets in a comprehensive way. The design also relates to the final task/application and for instance to the number of possible commands.

Designing an AR-BCI application often requires:

1. Testing hardware compatibility, depending on the EEG cap, the AR device and the exploited cortical area
2. Testing performance in real conditions (light, movement, etc)

Our experiments aimed to study (1) whether it is possible to combine a binocular OST-HMD with a BCI without any significant drop of performance in SSVEP recognition, and (2) whether it is possible to tolerate head movements and thus select moving targets in AR.

Our results showed that neither the wearing of the HMD, its electrical activity nor small head movements notably disrupted the BCI recognition accuracy. Thus, theoretically allowing us to use SSVEP in order to design user interfaces in AR.

Nonetheless, in order to make BCI-based AR systems usable, it is not enough to improve the hardware and software components of the system. The user, and particularly their training, should also be considered. Thus, investigating the user experience, for instance through the development of new feedback modalities to keep the user engaged and to help them modulate their brain activity in AR, seems necessary, as well as to evaluate the extent to which users

feel immersed and in control when interacting with a BCI/OST-HMD system. In this scope, it would be interesting to compare the combination of OST-HMDs and BCIs with other interaction technologies, both in terms of user experience and system performance.

4.4 Conclusion

In this Chapter we studied the combination of Brain-Computer Interfaces based on the SSVEP paradigm and Augmented Reality. First, we tested the feasibility of combining AR and SSVEP-based BCI technologies and assessed the influence of external and internal noise on the BCI accuracy within two user studies. We found that using an AR headset did not significantly impair the BCI performance and did not disrupt the EEG classification accuracy. Thus it seems possible to exploit the SSVEP and EEG data with an AR headset upon the EEG cap. We also found that it is possible to use an SSVEP-based BCI in presence of small head movements.

In this Chapter, our studies focused on the use of SSVEP for direct interaction in AR. In the following Chapter, we will study the possibility to use another BCI paradigm in AR: *Error Related Potentials*. In particular, we will investigate the presence of ErrPs in users facing different types of errors, often generated by AR systems.

CHAPTER 5

ERRPS: TOWARDS THE DETECTION OF SYSTEM ERRORS IN AR USING ERRPS

Abstract:

This Chapter aims at studying the presence and the detection of Error Related Potentials (ErrPs) in Augmented Reality. In particular, we investigate the presence and the detection of ErrPs when users face system errors, i.e errors originating from the system namely: tracking errors, feedback errors and background anomalies. The objective is to exploit the robustness and the controllability of Virtual Reality (VR) environments to simulate AR errors, in order to conduct the experiments in better controlled conditions. After describing the designed experimental protocol, we describe the materials and methods of the experiment before presenting and discussing the results.

5.1 Introduction

As seen before in *Chapter 1* an ErrP is elicited when the user commits or witnesses an error. It can be observed in the user's EEG signals as a specific EEG pattern appearing shortly after an erroneous event onset. The concept of erroneous event is relative to the user's perspective, and the amplitude of the pattern has been shown to be correlated with the gravity or impact of the error as well as with the engagement of the user in the task [Falkenstein et al., 2000]. Nevertheless, studies [Ferrez and Millán, 2008a] have shown that it was stable across time for each subject, and that it was possible to detect ErrP in single trials.

Error-related potentials can be particularly interesting in human-computer interaction (HCI) contexts [Chavarriaga et al., 2014]. Being able to detect when a user perceives an error can be useful in many scenarios. For example, it could be used to automatically assess the robustness of a system (e.g. linking the state of the system with erroneous events) or it could be used directly in the interaction process (e.g. automatically correct application or user errors).

Most of the existing works on ErrPs involve simple, unrealistic scenarios and do not study

ErrPs in an ecological environment (realistic environment, involving commonly found tasks) [Chavarriaga et al., 2014, Ferrez and Millán, 2008a, Spüler and Niethammer, 2015b, Omedes et al., 2015, Lopes-Dias et al., 2018]. In particular, the presence of and the possibility to detect ErrPs in Virtual or Augmented Reality has not been deeply investigated.

The goal of this Chapter is to study the possibility to use the ErrP paradigm in AR especially as AR often faces different sorts of errors among which, tracking and rendering errors. These errors can sometimes be difficult to automatically detect. Automatically detecting the presence of an ErrP when such an error happens, could potentially help automatically correct their effects.

As a first step towards this goal, we propose to assess the presence of ErrPs in users facing the previously mentioned errors in Virtual Reality (VR) environments. This choice is justified first, by the fact that VR environments are easier to control (as all the displayed elements are generated by the system) leading to more controlled experiments. Then, as the errors to which the users are submitted are the same as in AR, our hypothesis is that the findings will remain the same when transposed in AR.

We identify three common types of error in Virtual/Augmented Reality environments:

1. Tracking Errors: The tracking of an object being manipulated by the user is lost.
2. Feedback Errors: The system's response does not match the user's expectations.
3. Background Anomalies: An unrealistic event happens in the background

The possibility to detect the occurrence of these events from the users' brain activity paves the way to new means of improving interaction in Virtual/Augmented Reality. For the users, it could allow the design of mechanisms to automatically correct the system's behavior when they perceive an error, and consequently increase the systems' robustness. It could also help the systems' designers by providing them with a tool to better diagnose and review the interaction scenarios. One could think about automatically flagging the moments when users perceive errors, and marking them as points of interests.

In order to assess the presence and the possibility to detect ErrPs in the above mentioned situations, we designed a user study with the aim to highlight the neurophysiological pattern arising from: (1) Tracking errors, which prevent the user from accomplishing a task (2) Feedback errors, corresponding to the situation where the obtained feedback is erroneous; and (3) Background anomalies, corresponding to the situation where an unrealistic event, unrelated to the task, happens in the background of the VR scene.

5.2 Related Work

The first studies reporting a neurophysiological response to errors, were conducted in the early 90's [Falkenstein et al., 1991, Gehring et al., 1993]. They demonstrated that shortly after subjects committed errors in a speed response choice task, a characteristic EEG event-related potential (ERP) was elicited. This ERP was designated as *error-related negativity* (ERN). This ERN mainly consisted in a negative potential deflection over the fronto-central scalp areas appearing 50 to 100ms after a subject's erroneous response, and was followed by a centro-parietal positive deflection [Falkenstein et al., 2000].

Later studies further showed that the amplitude of the ERN was modulated by the importance of the error with regards to the given task [Frank et al., 2005], as well as with the subjective awareness of the error [Falkenstein et al., 2000, Navarro-Cebrian et al., 2013, Wessel, 2012]. In addition to the ERN, a similar EEG pattern was shown to appear after an erroneous feedback. Called the *feedback-related negativity* (FRN) this potential appears between 200 and 300 ms after the feedback onset [Hajcak et al., 2005]. Today, these patterns are grouped under the term *error-related potential* (ErrP) which has become widespread in the HCI community [Chavarriaga et al., 2014].

In the literature, at least 4 different types of ErrPs have been introduced, depending on the circumstances in which they occur. namely: (1) Execution errors, which correspond to errors made by the user during the execution of a task ; (2) Outcome error, which corresponds to the situation where the outcome of an action differs from what was expected ; (3) Observation errors, which correspond to the observation of another person or agent committing an error. (4) Interaction errors, which correspond to the execution of an unintended action by a HCI. Here, we study the errors originating from the interactive system, which fall in the "interaction errors" category.

5.2.1 Error-Related Potentials for Human-Computer Interaction

Originally, error-related potentials have been described in the context of BCIs by Schalk and colleagues [Schalk et al., 2000]. Error-related potentials can be used as a corrective signal, to correct a BCI's output when this does not provide the intended command [Ferrez and Millán, 2008b, Salazar-Gomez et al., 2017] or in an adaptive manner, reducing the possibility of future errors [Llera et al., 2011, Llera et al., 2012].

Additionally, the existence and decoding of ErrPs in asynchronous scenarios has been established [Kreilinger et al., 2012, Spüler and Niethammer, 2015b, Omedes et al., 2015, Lopes-Dias

et al., 2018]. Interested readers may refer to [Chavarriaga et al., 2014] for a more comprehensive review on error-related potentials.

These studies, suggest the possibility to detect and exploit ErrPs in interaction scenarios. However, most of them were conducted in non-ecological scenarios conditions. Recently, ErrPs have been studied in real-world situations, like driving a car [Zhang et al., 2015], indicating the feasibility of decoding ErrPs in ecological scenarios.

5.2.2 Error-Related Potentials in Virtual Reality

[Padrao et al., 2016] were among the first to investigate the presence of error-related potentials in VR. Their objective was to study the neurophysiological correlates of violating agency, and the differences between self-generated and external errors. Their experiment consisted in a fast selection task, involving arm movement towards the left or the right. The self-generated errors corresponded to errors made by the subjects by choosing the wrong direction, while the externally generated errors corresponded to a wrong feedback from the avatar, with the virtual arm going to a different direction from the subject's arm. Their results on 24 subjects show a clear difference between the neurophysiological responses corresponding to the 2 conditions. While the self-generated errors displayed a fronto-central negativity around 100 ms after the error onset, the errors generated by the avatar elicited a parietal negativity around 400 ms after the error onset.

Yazmir et al. [Yazmir et al., 2016] investigated the presence of error-related potentials in the brain activity of subjects submitted to errors in a hapto-visual task. The participants were asked to move a virtual object horizontally from an origin to a destination using a Phantom haptic arm. The path was obstructed by a large cylinder so that the participants would lose sight of the target behind the cylinder. The errors were generated by tweaking the vertical position of the target when reappearing from behind the cylinder. In the correct trials, the object would reappear at the same vertical level, while it would be randomly translated vertically in the error trials. Their results on 5 subjects suggest the presence of an ERN shortly after the error onset.

Later, in [Yazmir and Reiner, 2017] Yazmir et al. studied and compared the EEG correlates for success and failure in a "tennis-like" throwing game. The participants had to play the game against the computer. They considered 2 types of events: (1) Success or "hit", corresponding to the situation where the players scored a goal, meaning that they hit the ball and that the computer missed it. (2) Failure or "miss", corresponding to the situation where the players did not score, i.e. the computer repelled the ball. Their results show that negative peaks appeared in both of the events with different amplitudes and latencies.

In 2018, Pezzetta et al. [Pezzetta et al., 2018] investigated the presence of error-related potentials in a situation where the errors were more frequent than correct trials. Their goal was to disentangle and assess that ERN were indeed associated with a perceived error rather than with the rarity of the errors. They combined EEG measurement with a CAVE (Cave Virtual Environment) in an observation task, where participants had to watch a 1-st person perspective avatar realizing a grasping task. In the correct trials, the virtual arm would succeed in grasping a virtual glass, whereas in the erroneous trials it would miss the glass. The originality of this work, lies in the fact that the proportion of errors was higher than the correct trials (70% and 30% respectively). Nevertheless, their results on 25 subjects show the eliciting of a fronto-central negativity around 300 ms after the error onset, demonstrating the presence of observation errors in VR as well as the fact that despite their higher probability, erroneous events still elicited error-related potentials. More recently, [Gehrke et al., 2019] explored the use of ErrPs in order to detect conflicts in visuo-haptic integration. Their results suggest a more pronounced early negativity during the trials where the visual and haptic feedback did not match.

Previous works clearly suggest the presence of error-related potentials in the brain activity of subjects facing errors in VR. However, none of them systematically studied the different kinds of errors that can emerge from a VR system.

5.3 Materials and Methods

Our goal is to investigate the presence and detection of error-related potentials in different erroneous situations happening in VR. In particular, we identified 3 types of situations that potentially elicit error-related potentials: (Te) The loss of tracking of a manipulated object ; (Fe) An unexpected or erroneous feedback ; (Be) An unrealistic background anomaly. To assess whether or not these conditions elicit error-related potentials, we conducted a user study to record EEG activity under each one of them. From the literature and previous work, our hypotheses were that (Te) would elicit Event-Related Negativity (ERN) associated with execution errors, that (Fe) would trigger a Feedback Related Negativity whereas (Be) would trigger an ERN for the users who notice the anomaly and that it is possible to classify and detect the elicited ErrPs in single trial.

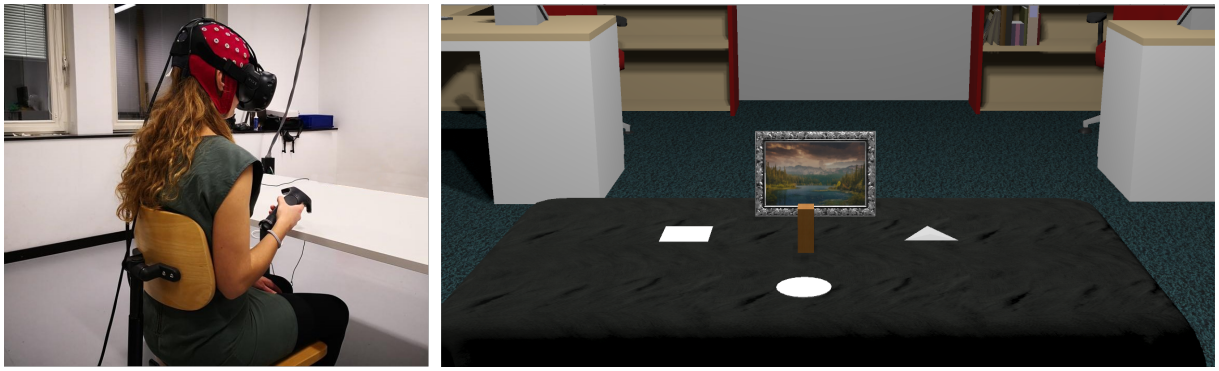


Figure 5.1: Experimental setup used in our study. Left : participant wearing both an EEG cap (ANTNeuro EEGO sports) and an HTC Vive VR headset. Right : virtual environment. Participants were seated in front of a virtual table on which three targets were displayed. The virtual object (brown) was placed in the middle of the three targets and had to be picked up and placed on one target.

5.3.1 Apparatus and participants

Fifteen healthy and not color-blind participants [1] (7 women, 8 men) took part in the experiment (aged mean= 24.8 years, std=2.9).

EEG data was recorded using an ANTNeuro EEGO Sports with 64 wet electrodes (AgCl Solution), grounded to the forehead, and referenced to CPz. The signal was amplified using an ANTNeuro Amplifier and recorded at 512 Hz.

The virtual scene and the experiment protocol were designed using the Unity software and C# scripting. The virtual scene was displayed using an HTC Vive Head-Mounted Display (HMD). The mounting of the HMD on top of the EEG cap was made in a way to avoid pressure on the fronto-central electrodes, by only fixating the HMD using the lateral elastic bands and leaving the upper elastic band looser.

The EEG data and experiment events were recorded using the Lab Recorder, which stores data sent using the Lab Streaming Layer (LSL) protocol [Kothe, 2014].

5.3.2 Experimental protocol

After the participants signed the consent form, they were equipped with the EEG cap and the VR headset. The virtual environment consisted of an office with a table placed in the middle,

1. This research was conducted in accordance with the relevant guidelines for ethical research according to the Declaration of Helsinki. All the participants were briefed about the nature of the experiment and signed an informed consent form at the beginning of the experiment.

associated with a real table to provide participants with a passive haptic feedback (see Figure 5.1). Participants sat at the table, and had to perform a center-out pick and place task, using a Vive controller for grabbing, moving and dropping the objects using the controller trigger.

The experiment consisted in 4 conditions corresponding to the types of trials:

- **Correct:** These trials correspond to the normal realization of the task. In these trials, the participant grabbed the object and dropped it in the correct shape. They were provided with the correct feedback (Fc) (Figure 5.2: Top left).
- **Tracking error (Te):** These trials correspond to the situation in which the system lost tracking of the object before the participant placed it into the target shape. At a random point within 25% and 75% of the distance between the initial position and the destination shape, the object froze and was detached from the participant's hand. In this event, the participant was asked not to go back to the object and just wait for the next trial (Figure 5.2: Top right).
- **Feedback error (Fe):** These trials correspond to the participant getting an unexpected, erroneous feedback after completing the task. Even if the object was placed in the correct shape, the shape would turn *red*, which corresponds to a wrong feedback (Figure 5.2: Bottom left). No further intructions were given to the participant.
- **Background anomaly (Be):** These trials correspond to the situation where an unrealistic anomaly, not related to the task, appears in the background. Randomly, when the object was within 25% and 75% of the total distance between the origin and the target shape, a frame placed in front of the participant on the table would flip and stay in an unrealistic position (crossing the table) until the end of the trial (see Figure 5.2: Bottom right).

Each trial started with the appearance of one of the three objects (a parallelepiped, a cylinder or a tetrahedron) placed at the center of the table. The participant's task was to grab the object and place it on the corresponding object's base shape (a square, a circle or a triangle respectively). After the participant placed the object on the corresponding shape, they were provided with a colored feedback for 2 s. The selected shape turned *green* if it was correctly associated with the object, and turned *red* otherwise. The timeline of the task is summarized in Figure 5.3.

The initial position of the object was equidistant to the 3 destination shapes (approx 25cm from the central position), and the whole setup could fit the user's field of view, so that the participant did not have to move the head to perform the task and thus to reduce muscle artifacts, which degrades EEG signals.

The experiment consisted of 10 runs of 39 trials (390 trials in total). In total there were approximately 69% of correct trials and 31% of erroneous trials. Each run consisted of 27 correct

Figure 5.2: Illustration of 4 experimental conditions. Top left: *Correct Feedback (Fc)* condition. The participant receives a correct feedback after the completion of the task. Bottom left: *Feedback error (Fe)* condition. The participant receives a wrong feedback after the completion of the task. Top right: *Tracking error (Te)* condition. The object freezes and is detached from the participant's hand. Bottom right: *Background anomaly (Be)*. The picture frame in front of the participant flips and gets into an unrealistic position, penetrating the table.

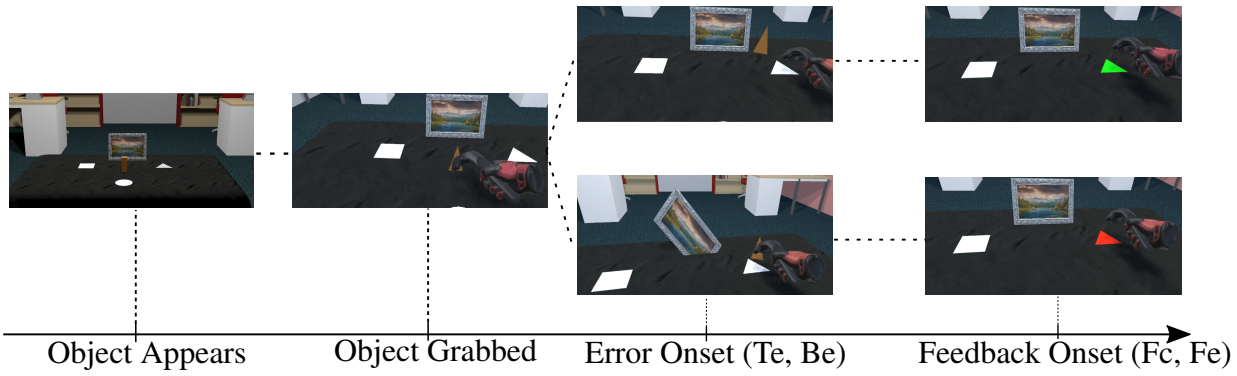


Figure 5.3: Timeline of a trial with the different possible events. At the beginning of the trial, the object appears at the center of the table and the participant had to grab the object. During the *tracking error (Te)* trials, the object freezes (top). During *background anomaly (Be)* trials, the frame on the tables flips (bottom). When the participant achieves the task, a correct feedback (top) is given in the *correct trials*, and an error feedback (bottom) is given on *Feedback error (Fe)* trials.

trials and 12 erroneous trials: 5 trials of condition Te (13%), 5 trials of condition Fe (13%) and 2 trials of condition Be (5%). For each run, the sequence of correct and error trials as well as the object in each trial were pseudo-randomly distributed, i.e., a random sequence was generated once and the same sequence was applied for all participants. The total duration of the experiment was around 1h30, including the time to set up the EEG cap, and, between runs, the participants could take breaks for as long as they wanted.

5.3.3 Electrophysiological analysis

Before being preprocessed, the signal was resampled to 256 Hz and re-referenced to the average of left and right mastoids (using electrodes P7 and P8). In order to perform artifact rejection, the signal was filtered between 1 and 30 Hz and only the channels F3, F1, Fz, F2, F4, FC3, FC1, FCz, FC2, FC4, C3, C1, Cz, C2, C4, CP3, CP1, CP2 and CP4 were considered. Epochs were rejected by threshold ($\pm$ 125 microvolts), probability and kurtosis. Regarding the methods based on probability and kurtosis, a threshold of 5 standard deviations was used as excluding criteria [Delorme et al., 2001]. For 14 participants (all except participant 10), up to 12.5% of trials per condition were removed. For participant 10, up to 25% of trials per condition were removed.

For the electrophysiological analysis, the data was band-pass filtered between 1 and 10 Hz with a zero-phase Butterworth filter of order 4.

In order to compare erroneous and correct responses, we defined three types of onset events in correct trials (Tc, Fc and Bc) occurring at comparable time-points to the error onsets in error trials (Te, Fe and Be, respectively).

For trials of type Fc, we considered as onset the moment in which the target shape turned green (correct feedback). Trials of type Tc and Bc had no intrinsic onset, so we considered virtual onsets corresponding to each participant's average delay of the corresponding error onsets (Te and Be, respectively), in relation to the moment in which the object was grabbed.

To display the eletrophysiological analysis of all conditions, we considered a 1 s window of starting -0.2 s before the corresponding onset and ending 0.8 s after it. The results are displayed as topoplots, calculated using all channels, and as EEG traces at channel FCz due to it being a representative channel.

5.3.4 Single Trial Classification

For the classification of the tracking condition (Te vs Tc), we considered 250 ms epochs starting 0.150 ms after either the error onset (Te) in the error trials or the virtual onset (Tc) in correct trials. For the classification of the feedback condition (Fe vs Fc), we considered 250 ms epochs starting 0.100 ms after either the error onset (Fe) in the error trials or the virtual onset (Fc) in correct trials. In both situations, the amplitudes of the channels FCz and Cz within these windows were used as features to train a shrinkage-linear discriminant analysis (LDA) classifier [Blankertz et al., 2011]. To address the imbalance between the classes, we considered, in each fold of the cross-validation, only a randomly selected subset of the correct trials, matching the number of error trials. This way, the classification was performed using a balanced data set.

For the classification of the tracking condition (Te vs Tc), Te was considered as the positive class and Tc as the negative class. Similarly, for the classification of the feedback condition (Fe vs Fc), Fe was defined as the positive class and Fc as the negative class. The classification results are reported in terms of accuracy, true positive rate (TPR) and true negative rate (TNR). The theoretical chance-level for all measures is 50%. The significance level (SL) was individually calculated for every participant, as the upper-limit of the 95% confidence interval of the theoretical chance-level, calculated based on the binomial cumulative distribution [Combrisson and Jerbi, 2015, Müller-Putz et al., 2008b].

5.4 Results

In the following we present the results regarding the electrophysiological analysis and the single-trial classification.

5.4.1 Electrophysiological Analysis

Tracking Error vs Correct Tracking Analysis (Te and Tc)

The results of the electrophysiological analysis of the tracking conditions (Te and Tc) are summarized in Figure 5.4. Figure 5.4a depicts the topoplots of the grand average correct (Tc) and error (Te) signals from $t = 0\,\mathrm{s}$ to $t = 0.55\,\mathrm{s}$ in intervals of 50 ms (top and bottom rows, respectively).

In figure 5.4b, the green line represents the grand average signal corresponding to correct trials (Tc) at channel FCz. The red line represents the grand average signal corresponding to

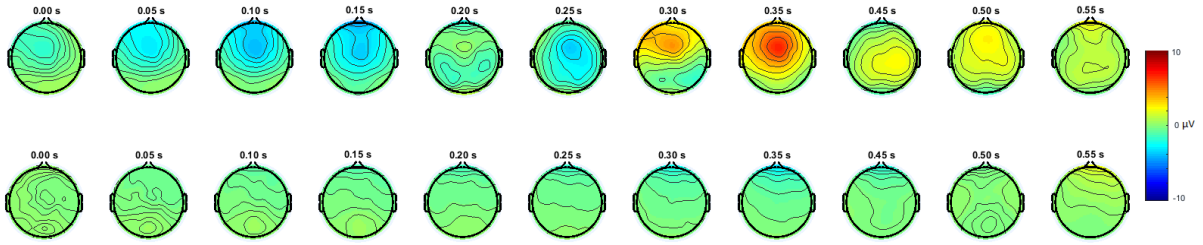


Figure 5.4a: Topoplots of the grand average error (Te) and correct (Tc) conditions (top and bottom rows respectively), displayed from $t = 0$ s to $t = 0.55$ s in intervals of 50 ms.

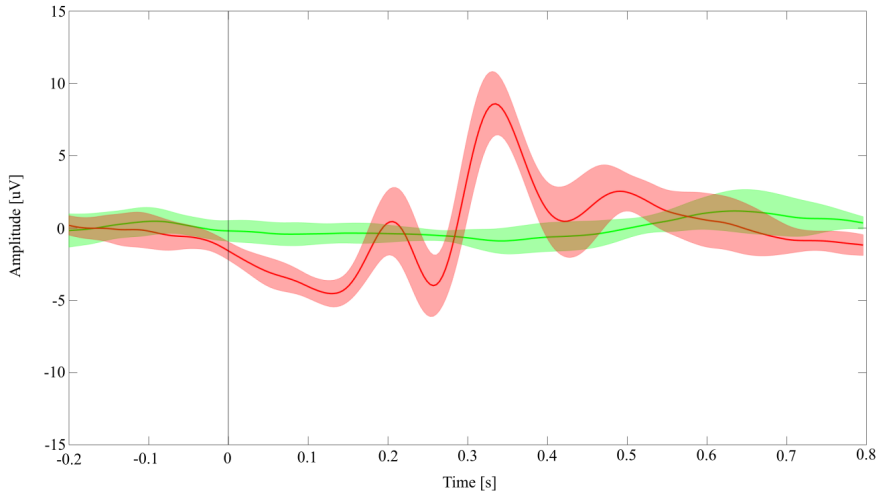


Figure 5.4b: Grand average correct (Tc) and error (Te) signals at channel FCz (green and red lines, respectively). The shaded areas represent the 95% confidence interval for the average curves. The black vertical line represents the onset events.

error trials (Te) at channel FCz. The correct epochs are not aligned to a real event but rather to a virtual onset, and therefore no event-related potential is expected. The grand average correct signal obtained is characterized by a relatively flat curve. The grand average error signal obtained is characterized by a positive peak at time $t = 0.206$ s with amplitude 0.4 µV, followed by a negative peak at time $t = 0.257$ s with amplitude -4.0 µV. This is followed by a positive peak at time $t = 0.335$ s with amplitude 8.6 µV, by a negative peak at $t = 0.421$ s with amplitude 0.4 µV and, finally, by a positive peak at time $t = 0.491$ s with an amplitude of 2.5 µV.

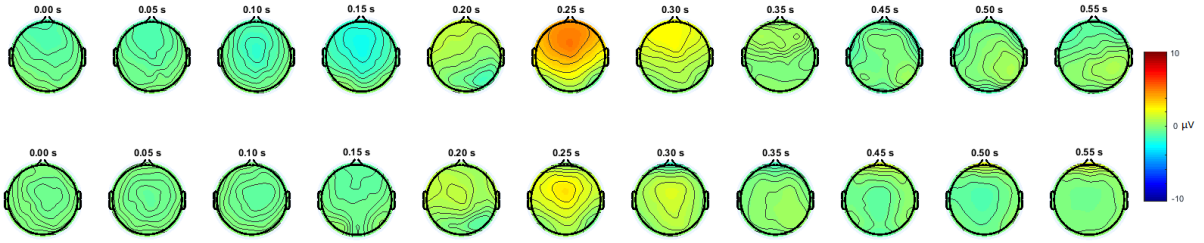


Figure 5.5a: Topoplots of the grand average error (Fe) and correct (Fc) conditions (top and bottom rows, respectively), displayed from $t = 0$ s to $t = 0.55$ s in intervals of 50 ms.

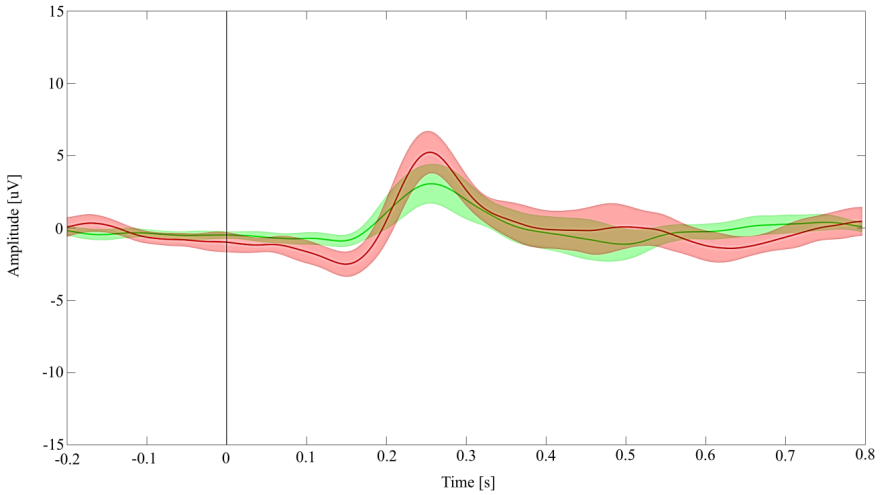


Figure 5.5b: Grand average correct (Fc) and error (Fe) signals at channel FCz (green and red lines, respectively). The shaded areas represent the 95% confidence interval for the average curves. The black vertical line represents the onset events.

Feedback Error vs Correct Feedback Analysis (Fe and Fc)

The results of the electrophysiological analysis of the feedback conditions (Fe and Fc) are shown in Figure 5.5. Figure 5.5a depicts the topoplots of the grand average correct (Fc) and error (Fe) signals from $t = 0$ s to $t = 0.55$ s in intervals of 50 ms (top and bottom rows, respectively).

In figure 5.5b, the green line represents the grand average signal corresponding to correct trials (Fc) at channel FCz. The red line represents the grand average signal corresponding to error trials (Fe) at channel FCz.

Both correct and erroneous feedback trials are time-locked to an event (start of feedback) and were expected to elicit an event-related potential.

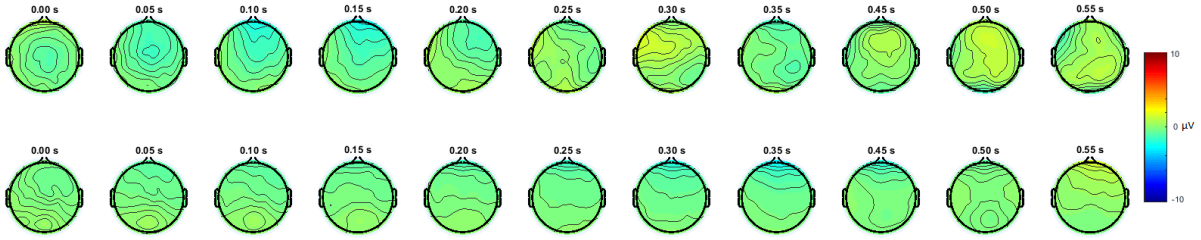


Figure 5.6a: Topoplots of the grand average error (Be) and correct (Bc) conditions (top and bottom rows, respectively), displayed from $t = 0$ s to $t = 0.55$ s in intervals of 50 ms.

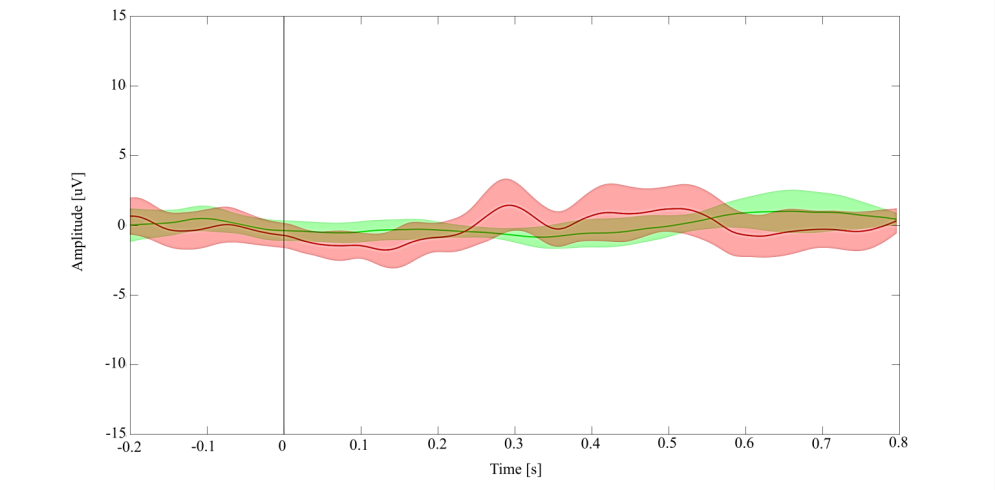


Figure 5.6b: Grand average correct (Bc) and error (Be) signals at channel FCz (green and red lines, respectively). The shaded areas represent the 95% confidence interval for the average curves. The black vertical line represents the onset events.

The grand average correct signal obtained is characterized by a positive peak at time $t = 0.257$ s with amplitude 3.05 µV. The grand average error signal obtained is characterized by a negative peak at time $t = 0.148$ s with amplitude -2.52 µV, followed by a positive peak at time $t = 0.253$ s with amplitude 5.22 µV.

Background Anomaly Trials Analysis (Be and Bc)

Figure 5.6 displays results of the electrophysiological analysis of the background anomaly conditions (Be and Bc). Figure 5.6a depicts the topoplots of the grand average correct (Fc) and error (Fe) signals from $t = 0$ s to $t = 0.55$ s in intervals of 50 ms (top and bottom rows, respectively).

In figure 5.6b, the green line represents the grand average signal corresponding to correct trials (Bc) at channel FCz. The red line represents the grand average signal corresponding to error trials (Be) at channel FCz. Correct epochs are not aligned to a real event but rather to a virtual onset and therefore were not expected to be associated with an event-related potential. Error trials were time-locked to the moment in which the frame flipped and were expected to cause an event-related potential, in case participants noticed the occurrence. The grand average correct signal obtained is characterized by a relatively flat curve. The grand average error signal obtained is also characterized by by a relatively flat curve, although a small positive peak at time $t = 0.292$ s with amplitude 1.43 µV is visible.

5.4.2 Single Trial Classification

In order to illustrate the feasibility of detecting error-related potentials in single trial in a VR experiment, we decided to test the single trial classification on the tracking condition (Te vs Tc) and on the feedback condition (Fe vs Fc). The EEG signals regarding the tracking condition were very distinct and Te condition presented a clear ErrP pattern. These were good indicators that these could lead to a good classification performance. The signals in the feedback condition were very similar to each other, posing challenges to a reliable classification.

Tracking condition

Table 5.1 displays the single trial classification results of Tc vs Te trials in terms of accuracy, TNR and TPR. The average accuracy obtained was 84.9%. The average TNR obtained was 88.4%. The average TPR obtained was 81.5%. The obtained results are significantly above chance-level ($\alpha = 0.05$).

Feedback condition

Table 5.2 displays the single trial classification results of Fc vs Fe trials in terms of accuracy, TNR and TPR. The average accuracy obtained was 59.3%. The average TNR obtained was 59.0%. The average TPR obtained was 59.6%. The average results are not significantly above chance-level ($\alpha = 0.05$).

Participants	Accuracy (%)	TNR (%)	TPR (%)	SL (%)
P1*	87.1 ± 7.5	90.0 ± 10.0	84.3 ± 10.4	60.2
P2*	91.8 ± 6.5	95.7 ± 6.2	88.0 ± 11.8	60.0
P3*	89.0 ± 6.3	88.9 ± 8.7	89.1 ± 10.2	60.3
P4*	91.7 ± 6.0	95.1 ± 6.8	88.2 ± 10.9	60.3
P5*	80.3 ± 9.6	81.3 ± 13.2	79.3 ± 12.9	60.3
P6*	81.6 ± 8.2	86.0 ± 13.2	77.3 ± 13.3	60.3
P7*	91.2 ± 6.0	94.6 ± 7.0	87.7 ± 11.8	60.1
P8*	73.8 ± 9.1	80.2 ± 11.5	67.3 ± 17.4	60.3
P9*	76.9 ± 11.5	80.2 ± 13.8	73.4 ± 15.7	60.2
P10*	81.9 ± 8.8	84.5 ± 11.5	79.3 ± 15.3	61.0
P11*	77.8 ± 9.7	79.6 ± 12.8	76.0 ± 14.8	60.3
P12*	79.9 ± 9.9	81.6 ± 13.0	78.4 ± 15.3	60.2
P13*	95.4 ± 4.4	98.7 ± 3.6	92.2 ± 7.5	60.1
P14*	85.1 ± 6.3	95.1 ± 6.9	75.2 ± 12.0	60.0
P15*	90.2 ± 5.6	94.4 ± 6.3	86.1 ± 9.8	60.1
Average	84.9 ± 6.5	88.4 ± 6.8	81.5± 7.1	

Table 5.1: Single-trial classification results of the tracking condition (Tc vs Te) in terms of accuracy, true negative rate (TNR) and true positive rate (TPR) for every participant (mean ± std) and their average. The Tc class was defined as the negative class and the Te class was defined as the positive class. The significance-level (SL) indicates the the upper-limit of the 95% confidence interval of the theoretical chance-level. Participants whose average measures were above significance level were marked with '*'.

Participants	Accuracy (%)	TNR (%)	TPR (%)	SL (%)
P1	56.7 ± 12.2	54.0 ± 16.3	59.3 ± 19.0	60.3
P2*	66.1 ± 9.3	65.2 ± 16.0	67.2 ± 12.6	60.2
P3*	68.1 ± 9.3	68.0 ± 14.1	68.2 ± 13.7	60.3
P4	54.4 ± 8.3	57.2 ± 16.3	51.7 ± 15.7	60.2
P5	60.5 ± 10.0	60.1 ± 16.0	61.2 ± 16.0	60.2
P6	62.8 ± 9.7	59.6 ± 17.4	66.2 ± 17.4	60.3
P7	54.5 ± 10.6	58.0 ± 13.2	51.0 ± 16.2	60.4
P8*	62.7 ± 9.9	63.1 ± 15.1	62.2 ± 17.1	60.3
P9	58.3 ± 10.6	58.3 ± 15.4	58.4 ± 18.7	60.6
P10*	64.2 ± 10.6	63.8 ± 15.6	64.8 ± 16.1	61.0
P11	48.1 ± 11.2	47.5 ± 17.8	48.8 ± 14.7	60.2
P12*	66.4 ± 9.5	60.3 ± 14.7	72.5 ± 15.5	60.2
P13	57.2 ± 10.5	57.4 ± 16.6	57.1 ± 17.7	60.2
P14	56.7 ± 10.8	61.3 ± 16.4	52.2 ± 18.8	60.2
P15	51.8 ± 9.3	51.1 ± 13.1	52.2 ± 18.9	60.2
Average	59.3 ± 5.8	59.0 ± 5.3	59.6 ± 7.3	

Table 5.2: Single-trial classification results of the feedback condition (Fc vs Fe) in terms of accuracy, true negative rate (TNR) and true positive rate (TPR) for every participant (mean ± std) and their average. The Fc class was defined as the negative class and the Fe class was defined as the positive class. The significance-level (SL) indicates the the upper-limit of the 95% confidence interval of the theoretical chance-level. Participants whose average measures were above significance level were marked with '*'.

5.5 Discussion

In this experiment we analyzed the electrophysiological signature during three possible system errors (tracking loss, feedback error and background anomaly) in a virtual reality task, and in particular the appearance of error-related potentials on the EEG data. A challenge in combining EEG measurements with VR is the fact that body and eye movements contaminate the signal quality.

Indeed, in our experiment, some participants' data was still contaminated with artifacts, even after artifacts rejection. For most participants, the tracking condition did not seem affected by artifacts, as participants were instructed to remain motionless when detecting a tracking error. No particular instructions were given regarding the feedback condition nor the background anomaly condition, making them more sensitive to artifacts. Additionally, the feedback condition was associated with the dropping of the object and the start of a new trial, making it prone to movement and blink artifacts.

The first observation that it was possible to make regarding the "tracking error" errors is that participants could easily notice them. As when they happened, participants were no longer able to achieve the pick-and-place task. Thus, this kind of error jeopardized the task and generated ErrPs. Besides, the pattern of the ErrP obtained in our study is very similar and well consistent with state-of-the-art literature regarding error signals in 2D paradigms [Spüler and Niethammer, 2015b, Omedes et al., 2014, Omedes et al., 2015].

The results of the analysis of the "feedback error" condition does not seem to show a clear ErrP pattern. Both correct and error trials were associated with an evoked potential. The erroneous condition displayed a small negative peak, that was not present in the correct condition, and displayed a positive peak with slightly higher amplitude than the correct condition. One possible explanation for this result is the fact that some participants stopped paying attention to the color of the given feedback. As the difficulty of the task was almost inexistent, participants did not rely on the feedback to evaluate their own performance. Hence, some of them may did not wait for the feedback before going back to the center of the scene and get ready for the next trial.

Lastly, our results showed that "Background anomalies" did not seem to generate any ErrP. One potential explanation could be that such kind of error had no impact at all on the execution or outcome of the task. The background anomaly could even remain unnoticed by some participants. Additionally, the low number of background anomaly trials can also hinder the results obtained (20 per participant, before artifact rejection). Thus, additional experiments need be conducted to further characterize the reactions to this specific kind of system error.

Overall, our results confirm previous results about the generation of ErrPs in Augmented or Virtual Reality. In our case, high-impact errors such as tracking errors generated ErrPs, while low-impact errors such as background anomalies did not generate any ErrPs. However, considering the ecological condition of the experiment and the fact that the participants were free of their movements, chances are significant that artifacts disrupted the ErrPs signals in some conditions, even after the artifact rejection procedure. Nevertheless, our results support the fact that high-impact errors could be detected in VR/AR using EEG analysis. In the context of VR/AR applications' assessment or debugging this would enable to automatically detect and grade the errors that a user might experience during a given session.

Moreover, the single-trial classification results for "feedback condition" were, on average, not significantly above chance-level. Even though 30% of the subjects were above. This could be a consequence of the ecological conditions of the study, leading some participants to stop paying attention to the color of the feedback.

The single-trial classification accuracy for the "tracking error" condition was 85% and thus significantly above chance-level, for all participants. This means a very good discrimination between classes "tracking error" and "no tracking error". Such high-level of detection performance is a starting point for asynchronous and real-time decoding of error-related potentials in VR/AR, as it was shown possible in 2D paradigms [Spüler and Niethammer, 2015b, Omedes et al., 2015, Lopes-Dias et al., 2018].

5.6 Conclusion

In this Chapter we have explored the possibility to detect error-related potentials in virtual reality when the system generates an unexpected error. In particular, we have investigated three types of errors: tracking errors, feedback errors and a background anomalies. The different errors represented a different aspect on the interaction process (task, outcome, context) and had a different impact on the realization of the task.

The results indicated that only the errors which had an impact on the task (tracking losses) were able to generate ErrPs since a clear ErrP pattern could only be observed in Figure 5.5b Additionally, we showed that the single trial detection of ErrPs related to tracking errors in VR was possible with a high accuracy. This is shown in Table 5.1 where tracking errors were detected significantly above chance level for all the subjects.

Taken together, these results show the viability of using BCIs to detect high-cost error-related potentials in a VR application which can lead to different use cases. In the future, this information

could be used for monitoring purposes, in order to detect application flaws, or in can be used for corrective purposes, in order to automatically correct errors.

In the next part, we will investigate the use of SSVEP for direct interaction in AR-BCI systems to propose new means of hands-free interaction in AR.

PART IV

Using SSVEP for interaction in AR

CHAPTER 6

IMPROVING SSVEP-BASED BCI: ASYNCHRONOUS CLASSIFICATION USING CCA AND HIDDEN-MARKOV MODELS

Abstract:

This Chapter aims at improving the self-paced classification accuracy of SSVEP responses, in order to improve the SSVEP-based interaction in AR. It introduces HCCA, as a new classification method for self-paced detection of SSVEP responses. HCCA is based on Hidden Markov Models (HMMs) that exploits the temporal dynamics of Canonical Correlation Coefficients (CC-Coefficients) to achieve asynchronous Steady-State Visual Evoked Potential (SSVEP) recognition.

6.1 Introduction

Steady-State Visual Evoked Potential (SSVEP) is one of the fastest EEG-based BCI [Chen et al., 2015]. It has been used for a wide variety of applications: Several BCI spellers have for example been developed by associating a flickering frequency to each key of a keyboard to allow typing on a computer screen [Hwang et al., 2012] or using a hierarchical approach by selecting groups of letters [Cecotti, 2010]. Other applications include the steering of robots [Si-Mohammed et al., 2018, Gergondet et al., 2011] or wheelchairs [Mandel et al., 2009].

Several methods have been proposed to detect SSVEP responses in EEG signals [Liu et al., 2014]. Historically, they have consisted in the analysis of the EEG power spectrum, looking for a significant increase of power around the stimulation frequencies [Cheng et al., 2002]. This method is one of the most intuitive ways to detect SSVEPs. However, the intensity of the

responses is associated to high within-and-between-subject variability [Müller-Putz et al., 2005], which can lead to poor recognition performances or to the necessity of frequent calibration phases.

In this Chapter, we propose to use Hidden Markov Models (HMM) [Rabiner, 1989] which are statistical models able to model and classify a sequence of feature vectors and thus exploit the temporal dynamics of these feature vectors. Using CCA coefficients as feature vectors, HMMs allow us to discriminate between no-control and control states as well as to perform SSVEP recognition. We called this method HCCA. In addition, HCCA was compared to a state-of-the-art method for self-paced SSVEP detection using EEG data recorded from 21 subjects, and was found to perform better, notably in terms of True Negative Rate (TNR).

6.2 Related Work

Recently, more efficient and more robust methods than the classical spectrum analysis method, have been proposed to detect SSVEPs [Lin et al., 2006, Cecotti, 2010, Friman et al., 2007]. In terms of feature extraction, we highlight 2 categories of methods: (1) Contrast methods and (2) Canonical Correlation (CC) based methods.

Contrast methods aim at increasing the Signal to Noise Ratio (SNR) of the recorded EEG [Friman et al., 2007]. The main idea is to consider an SSVEP response only when the signal significantly differs from noise. Volosyak et al. [Volosyak et al., 2010] and Sugiarto et al. [Sugiarto et al., 2009] applied Minimum Energy Combination to determine a spatial filter that would cancel as much noise as possible. To achieve this, all potential SSVEP components are removed through orthogonal projections. After this step, the obtained signal only contains nuisance signals and noise. The spatial filter is then determined so that it minimizes the energy of the resulting signal. Very similarly, Maximum Contrast Combination presented by Friman et al. [Friman et al., 2007], has been used by Wang et al. [Wang et al., 2010] to determine spatial filters that would maximize the ratio between the energy of the SSVEP frequencies and the energy of the nuisance signals. For further information about the different feature extraction methods –such as contrast methods for SSVEP recognition– interested readers may refer to Friman et al. [Friman et al., 2007].

Canonical Correlation Analysis (CCA) for SSVEP recognition was first introduced by Lin et al. [Lin et al., 2006]. Studies have shown its higher performance in terms of accuracy [Nan et al., 2011].

In synchronous designs, the classification of the canonical coefficients is usually done

following a maximum criterion [Lin et al., 2006]. The reference signal that is the most correlated with the EEG signal represents the selected frequency. However, for asynchronous interaction this maximum criterion is not enough because it does not allow to recognize no-control state.

[Horki et al., 2010] have proposed a thresholding approach to classify SSVEP responses. A response frequency is only considered if it is detected a certain percentage of times across a time window. If no frequency is above this threshold, no-selection is made. Another method based on thresholding of canonical correlation coefficients has been proposed by Xia et al. [Xia et al., 2013]. Called TRCC for Thresholding Ratios of Canonical Correlation, this method computes the ratio between the second and the first largest canonical correlation coefficients to discriminate between idle and work states. The user is considered in control state only if this ratio is below a certain threshold.

[Poryzala and Materka, 2014] proposed a method using canonical correlation coefficients as features to distinguish between idle and control states, named Clustering of CCA coefficients. Using the CCA Coefficients as feature vectors, they used training data to build centroids of the idle state class, and the control state class for each stimulation frequency using k-means clustering with the number of clusters set to 2 (for each stimulation frequency). Every new feature vector was then classified in each feature space, i.e., the pair of clusters of each stimulation frequency, based on its nearest centroid. In other words, a minimum distance to mean classifier is defined for each stimulation frequency, and a new input feature vector is classified as control or idle, relative to every stimulation frequency. If no classifier associates the feature vector to control state, then it is ultimately classified as idle. If only one classifier associates it to control state, then it is classified as a SSVEP response at the corresponding frequency. If more than one classifier associates it to control state, then a confusion leveraging step is performed as described in Section 6.4.3. Until now, this method was found to perform the best in terms of classification accuracy when compared to other contrast methods and CCA-based approaches [Poryzala and Materka, 2014, Suefusa and Tanaka, 2018].

The main observation that can be made from the above mentioned works is that they all rely on static approaches, i.e not taking into account the temporal evolution of the signal, and make decisions based on the canonical correlation values only at a certain time point, while it would be beneficial to exploit the dynamics and temporal evolution of the CC-coefficients.

6.3 The Proposed Method

HCCA is based on the hypothesis that when a user shifts their attention from no control or rest state to a flickering command, the CC-coefficient associated to the frequency of stimulation will tend to increase over time. Thus, exploiting this dynamic behaviour using HMMs would help improve the self-paced SSVEP recognition accuracy.

To summarize the approach, EEG trials corresponding to "no-control" or "SSVEP commands" are sliced into overlapping epochs. Each one of this epoch is described by its CC-Coefficients, i.e the canonical correlation of the epoch with reference signals corresponding to each one of the stimulation frequencies (see Section 1.5.3). This process results in a feature vector for each epoch of the trial. Concatenating the feature vectors corresponding to the CC-Coefficients of these successive epochs, describes the dynamics (evolution over time) of the CC-Coefficients.

The sequences of feature vectors of each class (corresponding to the "no-control" state and the stimulation frequencies) are then used to train HMMs. One HMM is trained for every class (see Section 1.5.4).

In a working mode, for every new EEG trial (sliced into epochs and described using CCA), the likelihood to have been generated by each one of the HMMs is estimated (see Section 1.5.4). The classification is then achieved following a maximum criterion. The class corresponding to the HMM with the largest likelihood, is designated as the class of the EEG trial. This whole approach is summarized in Figure 6.1.

6.4 User Study: Validation and Comparison of HCCA

In order to implement, test and compare HCCA on recorded data, we conducted a user experiment described in the following Section.

6.4.1 Apparatus and participants

Twenty-two naive participants (aged 27.6 ± 5.8 year-old, 6 women, 16 men) took part in the experiment [1]. Subjects had to perform a series of target selections consisting in focusing on a designated target from the 4 displayed targets (3 flickering targets and a fixation point). In order

1. The experiment was conducted in accordance with the relevant guidelines for ethical research according to the Declaration of Helsinki. All the participants were briefed about the nature of the experiment and signed an informed consent form at the beginning of the experiment. This study was approved by the Institution Ethics Committee (No 2019-22).

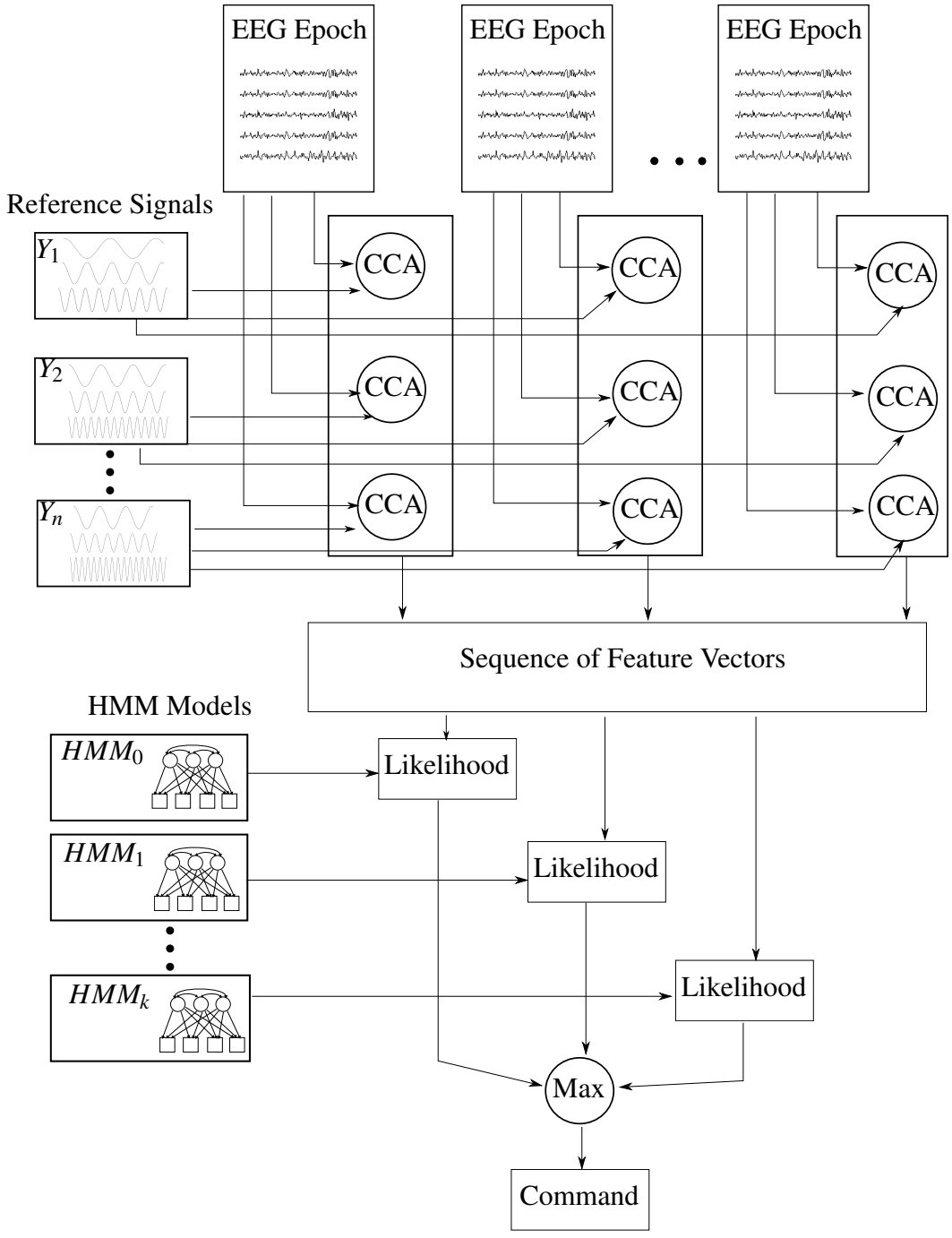


Figure 6.1: Global representation of the proposed method HCCA. After the training phase to estimate the HMM parameters, CCA is computed between successive epochs constituting a trial and every reference signal to determine the CC-Coefficients. The CC-Coefficients are aggregated in a sequence of feature vector representing the dynamics of the CC-Coefficients. The likelihood that each trained HMM model generated the sequence vector are computed and a maximum strategy is applied to determine the corresponding command.

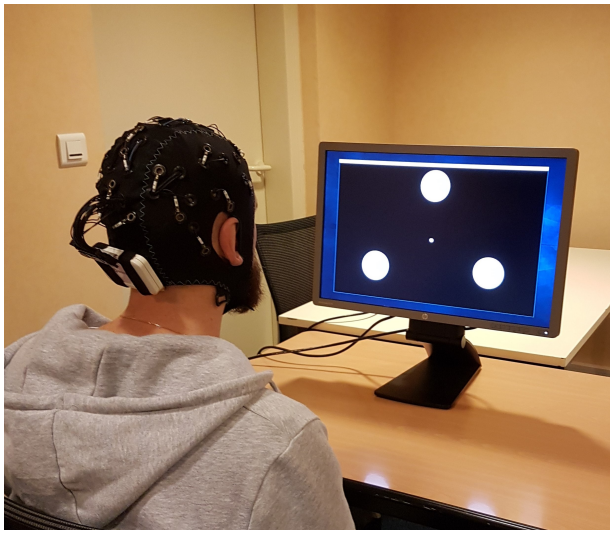


Figure 6.2: System Apparatus. The participant is wearing a 24-channels EEG cap and is sitting at 60cm from a LCD screen, where the 3 SSVEP targets are displayed.

to evaluate the ability of the proposed method to be used for self-paced SSVEP recognition, 3 of the targets were flickering while the central fixation point was not (see figure 6.2).

This latter was smaller than the targets and was used to represent the user's no control state. Rather than just rest state, this design allows to recognize the state where the users have stimulations in their visual field, but do not intend to send any command.

First, the participants were provided with all necessary information and signed the consent form. Then, they were equipped with the EEG cap (Figure 6.2). They sat on a chair at approximately 60cm from the computer screen displaying approx. 3cm wide white targets.

After that, participants were asked to complete 3 runs of SSVEP selections lasting 6min. 40sec each. Each run consisted in 40 trials (10 per target). Each trial consisted in focusing on one designated target following the scheme presented in Figure 6.3.

The target on which the participant had to focus during the coming trial was designated using a red pointing arrow for 1s before the 3 targets at the edges started flickering at 10, 12 and 15Hz for 5s, while the central fixation point remained still. The order of targets designation was pseudo-random. Following that, a 2s feedback was given by turning one of the targets green. As we did not perform on-line classification, the feedback was sham. More precisely, the designated target was highlighted in green 80% of the time (simulating a correct selection), while one of the 3 others was highlighted in red 20% of the time (simulating a mistake), in order to keep the participants motivated and engaged with the task. After 2 more seconds of break, a new target was designated and a new trial started.

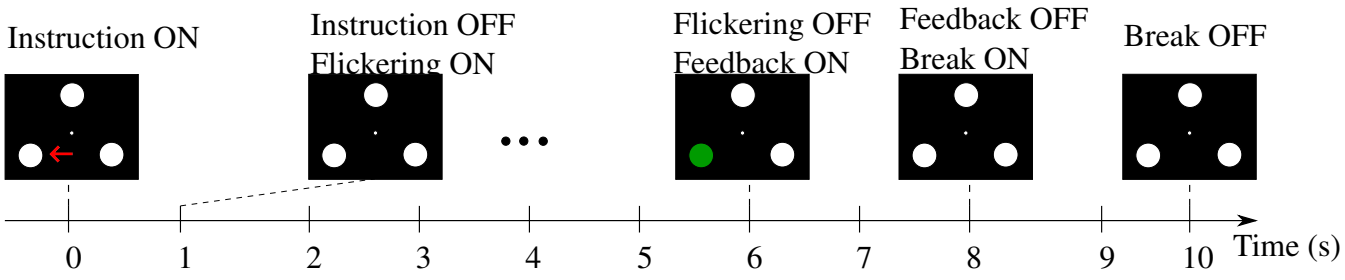


Figure 6.3: Timing of a trial. (t=0) The participant is instructed to focus on one of the displayed targets, with a red arrow. (t=1) The red arrow disappears and the 3 external targets start flickering for 5s while the central target remains still. (t=6) The flickering stops and and feedback is given to the participant (the designated target is highlighted in green 80% of the time). (t=8) A 2s-long break is given before the start of the next trial.

For each run, participants had to focus on each one of the targets 10 times (40 trials per run) presented in a pseudo-random order. In order to minimize visual fatigue, participants took as much break time as needed between the runs with a minimum break of 5 minutes.

Globally, the experiment enabled us to record 120 trials of SSVEP selections for an experiment duration of approximately 60 minutes.

6.4.2 Material and data acquisition

Visual stimuli were displayed on a 19-inch computer screen with a refresh rate of 60Hz (see figure 6.2). Electroencephalographic (EEG) data were recorded using a SMARTING amplifier (mBrainTrain, Serbia) with a sampling frequency of 500 Hz and using 5 scalp electrodes: O1, O2, POz, Pz and CPz referenced to Cz and grounded to AFz. The signal was recorded using the OpenViBE [2] Acquisition Server [Renard et al., 2010], and the sequence of trials as generated using Unity Software [3].

Following the recording, the signal was bandpass filtered between 5 and 40 Hz using a 4^{th} order Butterworth filter. After that, the 5s long trials were sliced into 2s long epochs with 1s of overlap (4 epochs per trial).

In order to describe the signal, we used the Canonical Correlation Analysis method keeping

2. OpenViBE v2.1.0
3. https://www.unity.com

only the largest CCA component per frequency to obtain a feature vector v_j

$$v_j = \begin{pmatrix} CC_{j_1} \\ CC_{j_2} \\ . \\ . \\ . \\ CC_{j_{N_f}} \end{pmatrix} \tag{6.1}$$

as described in Section 1.5.3

All training trials were turned into sequences of 4 CC-Coefficients feature vectors, corresponding to sequences of observation, and using the Expectation-Maximization (EM) algorithm [Rabiner, 1989], the parameters of the HMM (emission and transition probabilities) are estimated to best fit the obtained sequences as described in Section 1.5.4. We set the number of hidden states to 3.

The training phase results in a set of HMM corresponding to the no-selection state and each of the stimulation frequencies:

$$M = \begin{pmatrix} HMM_0 \\ HMM_1 \\ . \\ . \\ . \\ HMM_{N_f} \end{pmatrix} \tag{6.2}$$

Classification Phase

The classification phase consists in determining for a new sequence of feature vectors, which model from the previously trained HMMs best describes it, i.e., solving Problem 1. For each new sequence O, using the forward-backward procedure [Rabiner, 1989] we determine the likelihood $L_i(O)$ of the sequence being generated by the model i. The recognized class corresponds to the model that is the most likely to have generated the observed sequence. However, in order to reduce false positives, if the difference of the log-likelihood between the 2 HMM that are the most likely to have generated the observed sequence is below a threshold θ, the no-control state is chosen. This threshold was fixated to 4 in this study.

6.4.3 Clustering of CC-Coefficients (CCCA)

In order to assess the performance of HCCA method, we compared the results we obtained with the CCCA method proposed by [Poryzala and Materka, 2014]. The choice of CCCA is justified by the fact that this method, was found to be best performing self-paced SSVEP classification method in the literature [Suefusa and Tanaka, 2018, Poryzala and Materka, 2014].

CCA was performed between each 5s-long trial and each stimulation frequency. The 3 largest canonical correlation components were kept, as described in [Poryzala and Materka, 2014]. This process resulted in 3 feature vectors (one for each stimulation frequency). Each feature vector containing the 3 largest CCA components.

The order of the trials was randomized. One fifth of the trials was used for testing, while the rest of them was used for training following the procedure described in [Poryzala and Materka, 2014]. An iterative k-means clustering was performed using successive training trials to update the clusters' centroids of each feature space, until the distance between the centroids of each model exceeded the threshold β that had been set to 0.25, as described in [Poryzala and Materka, 2014].

For classification, every feature vector from the testing set was classified relatively to each stimulation frequency using minimum distance to each of the clusters' centroids. If the feature vectors was classified as no-control relatively to all the stimulation frequencies, it was globally classified as no-control. If the feature vector was classified as control, relatively to only 1 stimulation frequencies, the corresponding frequency was designated as the intended command. Lastly, if the feature vector was classified as control, relatively to more than 1 stimulation frequency, then a confusion leveraging step was performed. The Euclidean distance between the feature vector and the middle of the 2 cluster centroids was determined. The intended frequency was the one which the middle point between the clusters centroids was the largest [Poryzala and Materka, 2014].

All implementations were done using the *scikit-learn* Python library [Pedregosa et al., 2011].

6.4.4 Synchronous Evaluation

We first performed a synchronous (ignoring the no-control class) offline analysis of the recorded data. As subject 6 decided to stop the experiment before the end, their data was destroyed and the results presented thereafter do not include them. For each of the remaining 21 subjects, the trials corresponding to the 3 control states (10, 12 and 15Hz) were sliced into 2s-long epochs, with 1s of overlap. First, a CCA was performed on each of these epochs, and a

Max strategy was used to determine the recognized frequency for each epoch.

6.4.5 Self-Paced Performance Evaluation

Each 5s-long trial was sliced into 2s-long epochs, overlapping with 1s (4 epochs per trial). A CCA was performed between every epoch and each of the stimulation frequencies, resulting for each trial in a sequence of 4 feature vectors representing the dynamics of the CC-Coefficients over time.

We performed a 5-$fold$ cross-validation of the data recorded during the experiment, training on 4 fifth of the trials and testing on the remaining fifth.

In order to evaluate the performance of the considered methods, three performance measures were used (*True Positive Rate*, *True Negative Rate* and *Classification Accuracy*):

First, the True Positive Rate (TPR) and True Negative Rate (TNR) given by Equation 6.3 indicates the ability of the system to discriminate between selection and no-selection states:

$$\begin{cases} TPR = \frac{TP}{P} = \frac{TP}{TP+FN} \\ TNR = \frac{TN}{N} = \frac{TN}{FP+TN} \end{cases} \tag{6.3}$$

with TP and TN the number of *True Positives* and *True Negatives* respectively, and P and N the total number of *Positives* and *Negatives* respectively. A True Positive trial is defined as a selection trial correctly classified as such, and a True Negative is a no-selection trial correctly classified as such.

The third performance measure is the recognition accuracy, which corresponds to the ability of the system to discriminate between different SSVEP responses, ignoring the False Negatives and False Positives. It can also be viewed as the amount of correct commands the system sends.

6.5 Results

6.5.1 Synchronous Results

The synchronous analysis of the results allowed to distinguish between 3 groups of subjects from the 21 participants, which is consistent with the findings of Poryzala and Materka [Poryzala and Materka, 2014]:

- **Group A**: 9 subjects with over 90% of recognition accuracy.
- **Group B**: 6 subjects with a recognition accuracy between 60% and 90%.

Table 6.1: Recognition Accuracies per class after synchronous CCA analysis.

Subject	15Hz	12Hz	10Hz	Accuracy	Group
1	93%	89%	93%	91%	A
2	79%	79%	84%	81%	B
3	99%	98%	98%	98%	A
4	60%	73%	84%	72%	B
5	88%	99%	96%	94%	A
7	21%	52%	89%	54%	C
8	95%	100%	100%	98%	A
9	100%	100%	97%	99%	A
10	30%	51%	86%	55%	C
11	100%	98%	100%	99%	A
12	81%	72%	91%	81%	B
13	94%	98%	99%	97%	A
14	68%	77%	93%	79%	B
15	55%	66%	81%	67%	B
16	27%	25%	96%	49%	C
17	25%	52%	84%	53%	C
18	93%	97%	97%	95%	A
19	49%	61%	48%	52%	C
20	65%	71%	68%	68%	B
21	95%	100%	100%	98%	A
22	33%	64%	72%	56%	C

– **Group C**: 6 subjects with a recognition accuracy below 60% (chance level 40% [Müller-Putz et al., 2008a]).

The detailed results per subject are given in Table 6.1.

In a second time, in order to assess the temporal dynamics of the CCA coefficients, we plotted the mean coefficients corresponding to the successive epochs over time for all the subjects and for each class (see figure 6.4).

In line with our hypothesis, we noticed an increase in the CCA coefficient corresponding to the attended frequency.

Table 6.2: Confusion Matrix Using HCCA : Group A.

Group A	No-Control	15Hz	12Hz	10Hz
Idle	87%	2%	4%	7%
15Hz	5%	95%	0%	0%
12Hz	2%	0%	98%	0%
10Hz	7%	0%	0%	93%

Table 6.3: Confusion Matrix Using HCCA : Group B.

Group B	No-Control	15Hz	12Hz	10Hz
Idle	70%	8%	13%	9%
15Hz	24%	65%	3%	8%
12Hz	35%	2%	58%	5%
10Hz	31%	13%	2%	54%

Table 6.4: Confusion Matrix Using HCCA : Group C.

Group C	No-Control	15Hz	12Hz	10Hz
Idle	66%	10%	5%	19%
15Hz	28%	62%	2%	8%
12Hz	28%	7%	61%	4%
10Hz	42%	5%	7%	46%

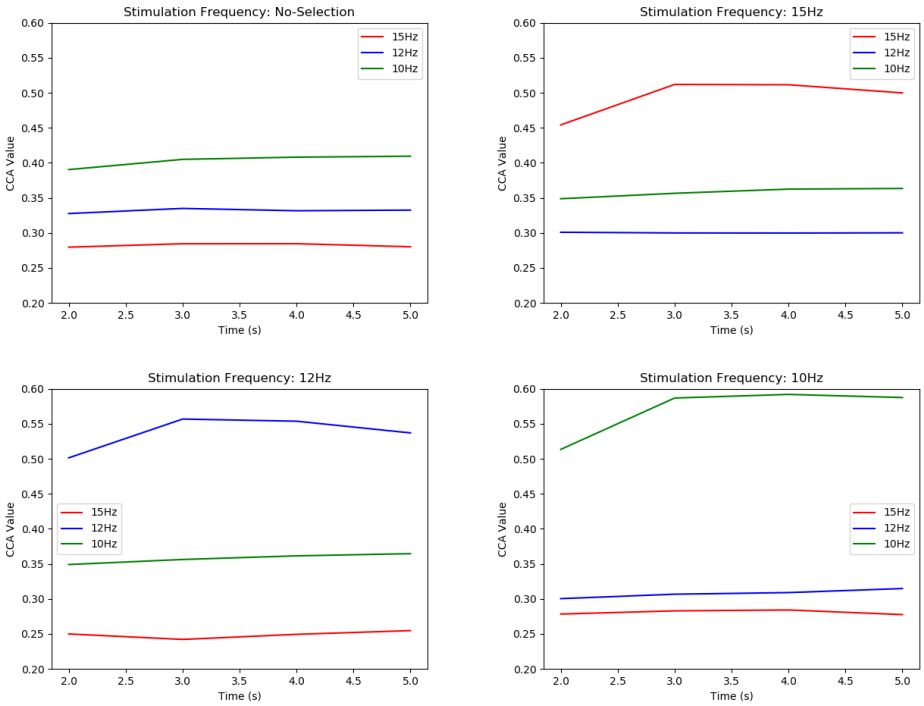


Figure 6.4: Illustration of the grand average evolution of CC-Coefficient over time (dynamics) for every state and stimulation condition. Each curve represents the largest CCA component of successive epochs with every stimulation frequency (Red: 15Hz, Blue: 12 Hz and Green: 10Hz). In the no-selection condition (top left) the 3 curves remain relatively still, while an increase of CC-Coefficient corresponding to the stimulation frequency is observed in the rest of the conditions. (2s corresponds to the end of the first epoch).

6.5.2 Self-paced Results

HCCA

The confusion matrices for groups A, B and C using HCCA are given in tables 6.2, 6.3,and 6.4 respectively.

The confusion matrices for groups A, B and C using CCCA are given in tables 6.5, 6.6 and 6.7 respectively.

Performance Comparison

Using the results from the cross-validation procedures described in the previous sections, we compared the results obtained with HCCA and CCCA in terms of TNR, TPR and Accuracy. The

Table 6.5: Confusion Matrix Using CCCA : Group A.

Group A	No-Control	15Hz	12Hz	10Hz
Idle	75%	5%	5%	15%
15Hz	14%	86%	0%	0%
12Hz	9%	0%	91%	0%
10Hz	7%	1%	0%	92%

Table 6.6: Confusion Matrix Using CCCA : Group B.

Group B	No-Control	15Hz	12Hz	10Hz
Idle	65%	3%	12%	20%
15Hz	20%	71%	3%	6%
12Hz	18%	2%	73%	7%
10Hz	20%	3%	1%	76%

Table 6.7: Confusion Matrix Using CCCA : Group C.

Group C	No-Control	15Hz	12Hz	10Hz
Idle	48%	8%	10%	34%
15Hz	21%	56%	3%	20%
12Hz	23%	5%	56%	16%
10Hz	25%	6%	6%	63%

results per subject and the means for groups A, B, C are summarized in tables 6.8, 6.9 and 6.10.

In order to test the significance of our results, we used a general linear mixed model (GLMM), the dependent variables were the TNR, the TPR and the Accuracy, the independent variable was the method (CCCA vs HCCA, within-subjects factor) and the user was modeled as a random factor. In addition, when the group had a significant effect on the results, it was introduced in the GLMM as a between-subjects factor. We tested the normality of the residuals using the Shaphiro-Wilk normality test and we applied the Greenhouse-Geisser correction when sphericity was violated. We used the Friedman ANOVA otherwise. The statistical analysis was performed using R-studio.

For the TNR, the Sphapiro-Wilk normality test could not reject the null hypothesis ($p = 0.302$), thus we performed a GLMM with the method (CCCA vs HCCA) as a within-subjects variable. The results showed a main effect on method ($F(1,20) = 6.94; p < 0.05; \eta_p^2 = 0.26$). Post-hoc tests (Bonferroni) confirmed that the difference was significant ($p < 0.05$). In overall, the HCCA method ($M = 76.00; SD = 14.37$) had a significantly higher TNR than the CCCA method ($M = 64.57; SD = 24.31$).

For the TPR, the normality of the residuals was also met ($p = 0.114$). However, as we observed an effect of the group, we added it to the model. Thus, we performed a GLMM analysis considering the group as a between-subjects factor (A,B,C) and the method as a within-subjects factor (CCCA, HCCA). As the GLMM showed an interaction effect between group and method ($F(2,18) = 10.0; p < 0.01; \eta_p^2 = 0.53$) we do not discuss the main effects. The post-hoc pairwise tests showed that for the group C, the method had a significant effect on the TPR. Participants of the C group had a higher TPR with the HCCA method ($M = 69.17; SD = 11.78$) compared to the CCCA method ($M = 53.17; SD = 2.48$). This effect was not observed for groups A and B.

Finally, for the accuracy, we did not find any significant effect on the group. However, as the residuals did not follow a normal distribution ($p < 0.05$) we used the Friedman ANOVA. The test did not show any significant effect for the method $\chi^2(1) = 2.67; p = 0.102$. In overall, both methods had a comparable accuracy rate $M = 90.19; SD = 13.02$.

6.6 Discussion

The results in terms of TPR and TNR confirm the effectiveness of using CC-coefficient dynamics with Hidden Markov Models for asynchronous SSVEP recognition. With average TNR of 87%, 65% and 66% for group A, B and C respectively, and an average TPR of 96%, 69% and 69% respectively, HCCA is able to accurately discriminate between selection and no-selection

Table 6.8: Comparison of the Recognition Accuracy per subject: Group A.

	TNR		TPR		Accuracy	
Subject	HCCA	CCCA	HCCA	CCCA	HCCA	CCCA
1	68%	76%	90%	84%	100%	100%
3	100%	80%	100%	92%	100%	99%
5	80%	20%	96%	83%	100%	100%
8	96%	96%	94%	94%	100%	100%
9	72%	64%	91%	91%	100%	99%
11	92%	60%	100%	92%	100%	100%
13	80%	96%	95%	90%	100%	100%
18	96%	84%	94%	89%	100%	100%
21	96%	100%	100%	95%	100%	100%
Mean	87%	75%	96%	90%	100%	99%

Table 6.9: Comparison of the Recognition Accuracy per subject: Group B.

	TNR		TPR		Accuracy	
Subject	HCCA	CCCA	HCCA	CCCA	HCCA	CCCA
2	80%	64%	72%	81%	100%	93%
4	68%	64%	58%	72%	54%	86%
12	68%	56%	74%	81%	99%	100%
14	72%	96%	77%	79%	80%	92%
15	68%	40%	70%	67%	85%	82%
20	64%	72%	65%	68%	80%	95%
Mean	65%	71%	69%	68%	83%	91%

Table 6.10: Comparison of the Recognition Accuracy per subject: Group C.

	TNR		TPR		Accuracy	
Subject	HCCA	CCCA	HCCA	CCCA	HCCA	CCCA
7	64%	48%	58%	54%	86%	75%
10	76%	36%	64%	55%	81%	76%
16	40%	16%	89%	49%	98%	83%
17	72%	44%	77%	53%	75%	68%
19	64%	64%	60%	52%	56%	63%
22	80%	80%	67%	56%	96%	87%
Mean	66%	48%	69%	53%	82%	75%

states. Moreover, with an average recognition accuracy of 100%, 83% and 82% for group A, B, C, the proposed method shows great performance in recognizing SSVEP commands, with only little confusion between commands.

The results of the statistical tests, showed a significant difference (p=.05) in terms of TNR between HCCA and CCCA, indicating that the proposed method better characterizes no-control states. However, no significant difference was found in terms of TPR and recognition accuracy, which suggests that the two methods are as effective as each other in terms of characterizing selection states and discriminating between SSVEP Commands. Overall, our results show that HCCA improves the self-paced SSVEP classification.

In a nutshell, these results confirm the potential held by dynamic modelling for characterizing and classifiying brain activity, despite their rather limited use in the literature [Lotte et al., 2007]. In particular, Hidden Markov Models seem to be particularly suited for self-paced BCIs. In fact, brain activity changes between ”hidden” mental states, that can only be observed through the measurement of particular phenomena and features.

This work is the first to successfully use Hidden Markov Models and Canonical Correlation Analysis for self-paced SSVEP recognition. In the future, we hypothesize that this technique could be adapted for online classification of SSVEP commands.

6.7 Conclusion

We presented here, a new method for asynchronous SSVEP recognition called HCCA. HCCA uses the dynamics CCA over time, when users shift their attention towards an SSVEP stimulation, through Hidden Markov Models which enable to model and classify these dynamics. Canonical Correlation Analysis was used for feature extraction, and the first canonical correlation components with the 3 flickering frequencies were aggregated as feature vectors. The sequence of feature vectors, representing the dynamics of the CC-Coefficients over time, were classified using Hidden Markov Models. One Hidden Markov Model was trained to characterize every selection class (every stimulation frequency), as well as the no-control state. Every new sequence of feature vectors was then classified according to the HMM that was the most likely to have generated it.

The user study conducted in order to evaluate HCCA and compare it to the state of the art, involved four conditions of SSVEP: 3 flickering targets and a central fixation point. This latter was meant to represent the no-selection state, where a user would have flickering targets in his field of view, without willing to send a command.

The results of the study showed the effectiveness of HCCA for asynchronous detection of SSVEP commands. This approach was evaluated and compared to Cluster Analysis of CC-Coefficients, which was previously found to be the best method in the state of the art, and statistical tests showed that HCCA better characterizes selection and no-selection states, holding a better True Negative Rate.

Taken together, these results suggest that HCCA is a promising technique for self paced interaction using SSVEP in AR. Provided that it is adapted to work online.

CHAPTER 7

DESIGNING AR INTERFACES FOR SSVEP-BASED BCIS

Abstract:

This Chapter aims at characterizing the design space of display strategies of SSVEP user interfaces in AR. It first introduces a design space of the different display strategies of a 3-commands SSVEP system to steer a mobile robot on a horizontal plane. We identified 5 dimensions namely: Orientation, Frame-of-Reference, Anchorage, Size and Explicitness. We evaluated the users' perception of intuitiveness regarding each display strategy. Then, we evaluated the effect of these dimensions on the classification accuracy of SSVEP selections.

7.1 Introduction

When designing interactive systems in AR based on BCIs, in addition to the classification performance, the User Interface plays an important role for the usability of the system. In the case of SSVEP, the UI lies in the layout of a selection menu and the strategy to display several targets (typically 3) to the user. This layout however, is strongly related to the task and the application that it is designed for. In the case of this Chapter, we chose the illustrative example of a robot teleoperation system.

The objective of this Chapter is to propose a generic methodology to design User Interfaces for BCIs based on SSVEP. We considered, as a starting use-case, the context of robot control, using 3 commands. We present an extended design space of different possible layouts, to display 3 targets for SSVEP-based robot control. This design space comprises five dimensions of 2 modalities: *orientation* (frontal vs. transversal), *frame-of-reference* (ego-centered vs. exo-centered), *anchorage* (user-attached vs. robot-attached), *size* (absolute vs. adaptive) and *explicitness* (explicit vs. implicit). This design space implies 32 (2^5 combinations) different display strategies. We then conducted a user study to select the preferred strategies among 32 which enables to rank them and identify the most intuitive candidates. Following that, we evaluated 4 representative

strategies, in terms of BCI performance on a concrete case of robot control.

To illustrate the development of a prototype of BCI-based application in AR. We could integrate all our previous results into a unique setup dedicated to the control of a real mobile robot in AR by means of an SSVEP-based BCI.

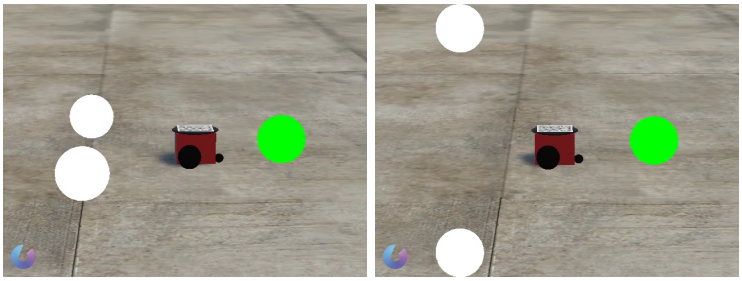
7.2 Design Space of SSVEP Target Display Strategies

Our objective is here to propose a design space that would characterize all the possibilities to display 3 SSVEP targets in the 3D space. Considering our application case of mobile robot control, three commands are enough for steering the robot with 3 options: “Forward", “Turn Left" and “Turn Right". In order to propose the most coherent association of each command with the AR target, we propose a 5-dimension design space that describes all possible layouts of 3 command targets. These 5 dimensions, namely *Orientation, Frame-of-reference, Anchorage, Size and Explicitness*, are presented beneath:

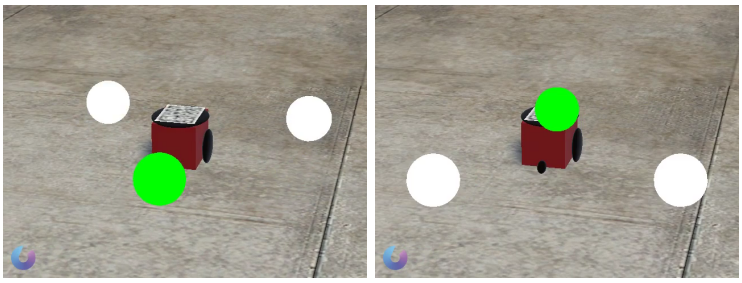
- **Orientation**: This dimension corresponds to the orientation of the plane containing the 3 targets. This plane can be “Transversal" i.e. parallel to the horizon or “Frontal" i.e. facing the user (see Figure 8.18a). When arranged transversally, the targets are in the same plane as the robot is moving. We hence expect that it would be easier to associate every command with its corresponding effect. When arranged frontally, the user has to make a mental rotation along the horizontal axis.
- **Frame-of-reference**: This dimension corresponds to the frame of reference of the 3 targets. It can be set to the robot, or to the user. It can be associated to the concepts of “ego-centered" and “exo-centered" frames of reference. In the first case, the targets correspond to absolute direction, set relatively to the user. For associating a target with a direction command, the user has to perform a mental rotation to map the targets with the robot direction of movement. When “exo-centered" i.e. attached to the robot, the plane containing the targets would rotate with the rotation of the robot and thus, keep the targets’ arrangement constant relatively to the robot. In this second setup, the target at the right side of the robot will always correspond to a right turn (see Figure 8.18b).
- **Anchorage**: This dimension corresponds to the position of the targets’ frame of reference in 3D. The targets can be either “anchored" to the robot or to the user. In the first case the three targets are set on the top of the robot and thus move with the robot. In the second case they are attached to the field of view of the user - at a distance of 2 meters to meet the focal distance configuration of the HMD. In such case, they would move accordingly to users’

head movements (see Figure 8.18c). The main difference between the two modalities lies in the fact that when anchored to the user, the latter can issue commands to the robot even when it is no longer in the user's field of view.

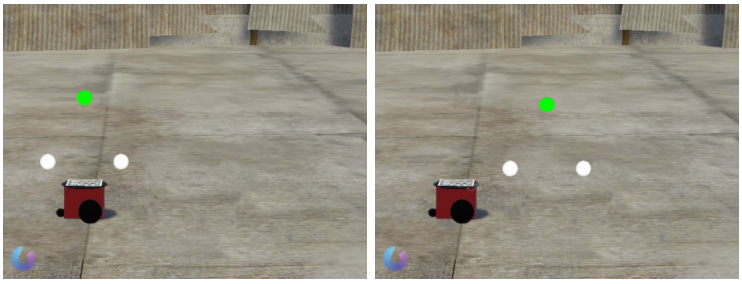
– **Size**: This dimension relates to the size of the targets, which can also be an influential parameter for the system ergonomics. The targets' size can be either "Absolute" or "Adaptive". As the targets are 3D objects, we used this terminology to define their real size. When the size in set to absolute, the radius of the target is fixated in the 3D space. In other words, the farther the robot is from the user, the smaller the targets appear. With the perspective, the size of the targets depends on the distance between the user and the robot. In the adaptive condition, the real size of the targets in the 3D space is not fixated. It would adapt to always be displayed at the same size relatively to the user. In other words, even if the robot is far from the user, the size of targets increases to compensate, and thus the targets always appear at the same size to the user (see Figure 8.18d). The advantage of the absolute setup lies in the fact that it is easier to associate the targets with the robot they control. However, bigger targets can be more comfortable to focus on compared to smaller ones. Lastly, it has been reported [Ng et al., 2012] that larger SSVEP targets elicit stronger SSVEP responses, we thus expect better classification results in the configurations where the size is set to adaptive.
– **Explicitness**: Rather than a display configuration, this dimension corresponds to the explicitness of the association between the targets and the commands. One way to improve the semantics of the application is to explicitly link a target with the direction it leads towards. When the explicitness is set to "Explicit", we chose to link the target with the corresponding wheel of the robot (front wheel, right wheel and left wheel) with a visual red string (see Figure 8.18e). When set to "Implicit", no red string was displayed. Even though, this additional content may overcharge the user's field of view, we hypothesize that it will help to seamlessly associate the target with the command.



(a) Orientation: Transversal (left), Frontal (right)

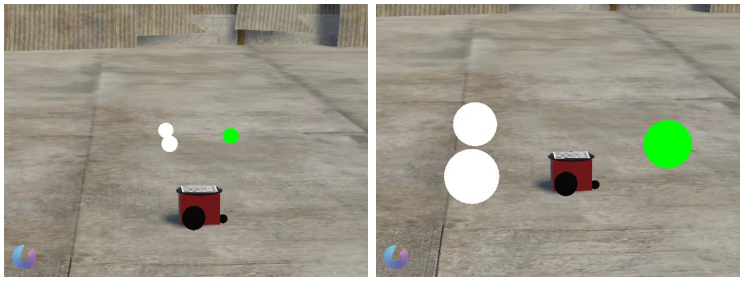


(b) Frame of Reference: Exo-Centered (left), Ego-Centered (right)

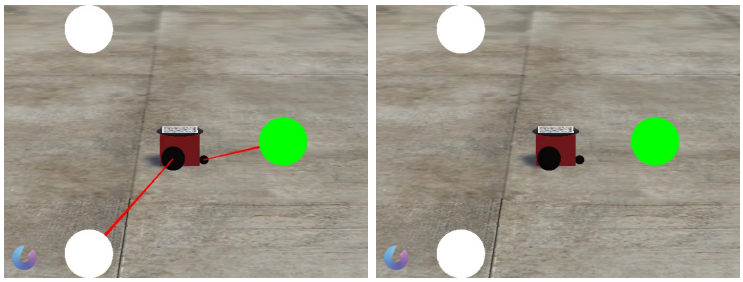


(c) Anchorage: Robot (left), User (Right)

Figure 7.1: Illustration of the 5 dimensions of our design space of SSVEP targets layout and display strategies. Each couple of figures represent the 2 modalities of 1 dimension: (8.18a) the *orientation* of the plane containing the targets, (8.18b) the *frame-of-reference* of the targets coordinates, (8.18c) the *anchorage* of the targets and their position



(a) Size: Absolute (left), Adaptive (Right)



(b) Explicitness: Explicit (Left), Implicit (Right)

Figure 7.2: (8.18d) the *size* of the targets which can be absolute or adaptive and (8.18e) the *explicitness* of the association target/command.

7.3 User Study 1: Subjective Preference of the Targets' Display Strategies

We conducted a user study in order to evaluate the different display strategies according to their subjective intuitiveness and coherence relatively to the task. Our goal was to determine the command layouts that the user would perceive as the easiest to understand considering the scenario of robot control.

7.3.1 Experimental Protocol

We recorded 32 video shots of 24 seconds each, corresponding to all the possible display strategies (reader may refer to accompanying video). All the 32 videos displayed the same virtual mobile robot executing the same path with different display strategies. This path was designed so to have the same number of command occurrences (4 turning right, 4 turning left and 4 moving forward) in addition to display the robot from the 4 possible angles (left to right, right to

left, back to front and front to back) for every strategy. Thus, no particular configuration was over-represented compared to the others, which would have potentially biased the results. Every time a command was issued to the robot, the corresponding target turned to green right before and during the execution of the command, in order to highlight the association between the command and the target.

An online and anonymized questionnaire filled by 42 participants, enabled to compare the different display strategies by asking the participants to evaluate “how coherent/intuitive was the association between the targets and the commands" for each video on a 7-point Likert scale (Reject, Poor, Acceptable, Satisfactory, Good, Very Good and Excellent). The 32 videos were displayed in a fully randomized order for each participant, in order to minimize ordering effects.

The final ranking of the strategies was performed following a majority judgment procedure [Balinski and Laraki, 2011]. This method consists in 3 steps. First, the majority-grade of an item (display strategy) is defined: it corresponds to the median grade voters attributed to this item. Second, the majority-grade should be completed by a + or - sign depending on whether more voters attributed a higher or lower grade to this item, respectively. Finally, the majority-ranking enable to arrange the items according to their majority-grades. In order to perform this majority-ranking, the majority-gauge of each item should be computed. This gauge is composed of three values (p, α, q), where α is the majority-grade and p/q are the percentage of grades above/below the majority-grade, respectively.

The videos were named from 0 to 31 according to the binary value of each dimension: $N = O*2^4 + F*2^3 + A*2^2 + S*2^1 + E*2^0$ with O,F,A,S,E (standing for Orientation, Frame-of-Reference, Anchor, Size and Explicitness dimension respectively). This notation enabled us to interpret the clustering obtained after the evaluation.

7.3.2 Results and Discussion

Table 7.1 represents the results of the majority judgment for each of the 32 videos and display strategies. As a result of this method of ranking, we could cluster the videos in 4 groups according to the values of 2 dimensions (“Orientation" and “Frame-of-reference") (see Figure 7.1). The group of videos with an exo-centered frame-of-reference and with a transversal orientation all have a “good" majority grade evaluation (videos from 25 to 32). The group with frontal orientation and an exo-centered frame-of-reference has a majority judgment of “satisfactory" while the two other groups have a “poor" or “acceptable" dominant majority grade evaluation.

Taken together, our results suggest strong trends regarding the subjective preference of the users. In a nutshell, we found that participants preferred a transversal orientation, meaning that the

targets should remain in the same plane as the robot's motion i.e. the horizontal plane here. More importantly, the user strongly preferred when the frame-of-reference for the targets coordinates was set to the robot (exo-centered). We hypothesize that these two configurations (transversal orientation and exo-centered frame-of-reference) minimize the mental rotation necessary to map the command with the direction. However, the highest subjective preference of the users does not guarantee the best objective performance of the SSVEP detection. Therefore, in the following section, we describe another user study meant to study the influence of the UI and display strategy on the BCI performance.

7.4 User Study 2: Influence of the Targets' Display Strategy on BCI Performance

To quantitatively evaluate the influence of the targets' display strategy on the classification performance, we conducted a user study in which we kept the best ranked and most representative display strategy (DS) from each of the 4 groups of strategies previously obtained (in section 7.3.2). We selected the best strategy from each group, provided that at least 2 dimensions differ between each 2 strategies. This method resulted in the selection of the strategies DS0, DS15, DS22 and DS29 corresponding to:

- DS0: ("Frontally-oriented", "Ego-centered", "User-anchored", "Adaptive-sized" and "Implicit") display strategy (see Figure 7.3a)
- DS15: ("Frontally-oriented", "Exo-centered", "Robot-anchored", "Absolute-sized" and "Explicit") display strategy (see Figure 7.3b)
- DS22: ("Transversally-oriented", "Ego-centered", "Robot-anchored", "Absolute-sized" and "Implicit") display strategy (see Figure 7.3c)
- DS29: ("Transversally-oriented", "Exo-centered", "Robot-anchored", "Adaptive-sized" and "Explicit") display strategy (see Figure 7.3d)

DS0 is considered here as the control condition, since this display strategy is similar to a standard SSVEP training phase and known to maximize the SSVEP response [Ng et al., 2012].

7.4.1 Apparatus and Participants

We implemented an augmented reality playground with a highlighted path that a virtual robot had to move through as illustrated in figure 7.4. This path was designed so that all of the three command directions are used the same amount of times to complete it.

Table 7.1: Results of user study 1 (N=42). The majority grade corresponds to the median, which means that at least 50% of the participants evaluated the strategies at least as the majority grade. Four groups, corresponding the 4 possible combinations of the Orientation and Frame-of-reference dimensions, emerge according to the dominant majority grade: poor, satisfactory, acceptable and good for [1-8], [9,16], [17-24] and [24-32] resp. (**O**: Orientation, **FoR**: Frame-of-reference, **A**: Anchorage, **S**: Size, **Exp**: Explicitness, **T**: Transversal, **F**: Frontal, **Ex**: Exo-centered, **Eg**: Ego-Centered, **R**: Robot, **U**: User, **E**: Explicit, **I**: Implicit.). The color coding of columns O and FoR refer to the 4 groups defined by the combinations of the 2 variables.

Ranking	Display Strategy #	Dimensions					p %	$\alpha\pm$	q %
		O	FoR	A	S	Exp	a	a	
1	29	T	Ex	R	Ad	E	48%	Good+	17%
2	30	T	Ex	R	Abs	I	45%	Good+	31%
3	31	T	Ex	R	Abs	E	43%	Good+	29%
4	28	T	Ex	R	Ad	I	40%	Good+	31%
5	26	T	Ex	U	Abs	I	31%	Good+	29%
6	24	T	Ex	U	Ad	I	33%	Good-	33%
7	27	T	Ex	U	Abs	E	31%	Good-	43%
8	25	T	Ex	U	Ad	E	21%	Good-	43%
9	15	F	Ex	R	Abs	E	33%	Good-	45%
10	8	F	Ex	U	Ad	I	26%	Good-	50%
11	9	F	Ex	U	Ad	E	48%	Satisfactory+	33%
12	10	F	Ex	U	Abs	I	45%	Satisfactory+	33%
13	11	F	Ex	U	Abs	E	43%	Satisfactory+	36%
14	14	F	Ex	R	Abs	I	38%	Satisfactory-	43%
15	12	F	Ex	R	Ad	I	36%	Satisfactory-	50%
16	0	F	Eg	U	Ad	I	31%	Satisfactory-	50%
17	13	F	Ex	R	Ad	E	48%	Acceptable+	31%
18	16	T	Eg	U	Ad	I	38%	Acceptable-	43%
19	22	T	Eg	R	Abs	I	33%	Acceptable-	43%
20	2	F	Eg	U	Abs	I	26%	Acceptable-	45%
21	1	F	Eg	U	Ad	E	38%	Acceptable-	50%
22	7	F	Eg	R	Abs	E	24%	Acceptable-	50%
23	4	F	Eg	R	Ad	I	45%	Poor+	24%
23	6	F	Eg	R	Abs	I	45%	Poor+	24%
25	20	T	Eg	R	Ad	I	45%	Poor+	26%
25	5	F	Eg	R	Ad	E	45%	Poor+	26%
27	17	T	Eg	U	Ad	E	45%	Poor+	36%
28	18	T	Eg	U	Abs	I	45%	Poor+	38%
29	3	F	Eg	U	Abs	E	40%	Poor+	38%
30	21	T	Eg	R	Ad	E	33%	Poor-	36%
31	23	T	Eg	R	Abs	E	36%	Poor-	38%
32	19	T	Eg	U	Abs	E	38%	Poor-	45%

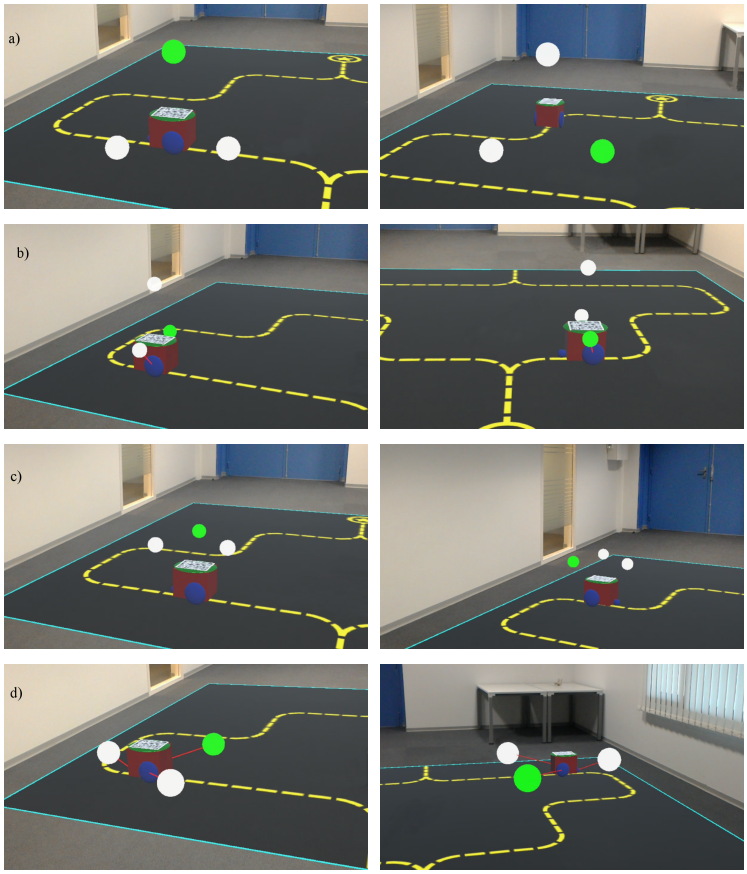


Figure 7.3: Illustration of the 4 display strategies used in user study 2. DS0 (a), DS15 (b), DS22 (c), DS29 (d).

The virtual playground and all its components (path, robot and command targets) were displayed on a Microsoft HoloLens. The EEG headset was the same as described in Section 4.2. Twenty-four participants (aged 28.1 ± 8.1, 4 women) took part to this experiment.

7.4.2 Experimental Protocol

After they signed the informed consent and were given all relevant information, the participants were equipped with both the EEG and the AR headsets. The experiment was composed of two parts: (1) a training acquisition phase that lasted 7 min and that was used to gather the data to train the classifiers, and (2) an evaluation phase that was used to compare the 4 display strategies in terms of BCI performance.

The training phase consisted of a unique run of 30 trials (SSVEP selections). The targets

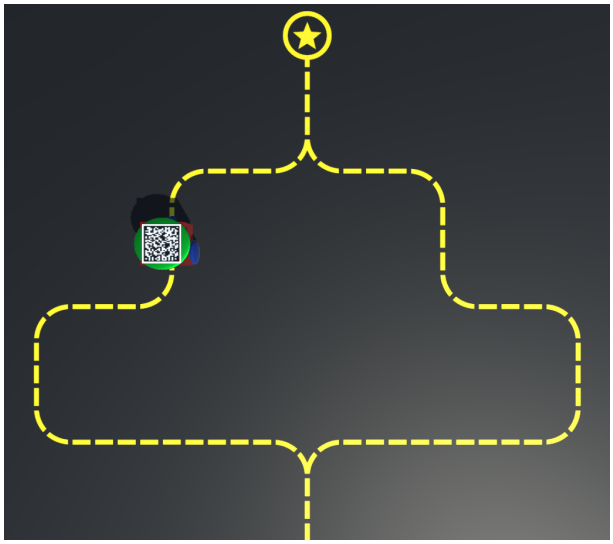


Figure 7.4: Virtual playground and highlighted path to follow.

were 10 cm wide circles arranged in an equilateral triangle of 46 cm side placed at a distance of 2 m from the user. They were flickering at 10, 12 and 15 Hz. The trial structure was the same as described in Figure 4.2.

Following the training phase, the evaluation phase consisted in 4 runs of 18 trials (4 min). Each run was using one of the selected display strategies, and consisted in a round-trip of 9 trials. Each trial consisted in focusing on the designated command target for 7 seconds, followed by 4 seconds of feedback (the recognized target), during which, the robot performed the given action. In order to maximize the consistency between the participants, they only made selections while the robot was immobile. Moreover, the robot always performed the correct move and the feedback was sham. The participants were informed that the robot would always follow the right path, but they were not told about the Sham aspect of the feedback. Hence, to keep a high level of engagement, their objective was to maximize the number of correct detections. In order to avoid any order effect that would bias the results, the order of the runs was counterbalanced across participants (24 possible arrangements for 24 participants).

7.4.3 Results and Discussion

Following the results of the literature [Ng et al., 2012], our hypothesis was that the participants would perform the best in DS0 condition as it optimizes the size of the targets and the distance between them. In order to compare the SSVEP recognition performance between all the

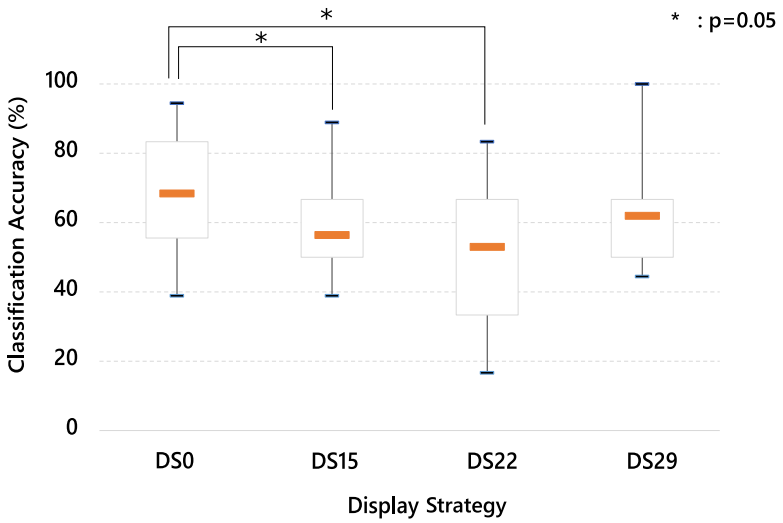


Figure 7.5: Results of User Study 2: Boxplots representing the classification accuracy, in %, (N=21) as a function of the display strategy.

conditions (DS0, DS15, DS22 and DS29), we used the data gathered during the training phase to train the CSP filters and the LDA classifiers for each participant. Every 7s trial was subdivided into 500 ms epochs, with an overlap of 400 ms. Each epoch was classified and majority voting across all the epochs determined the recognized class.

Our analysis considered the percentage of correct responses for the entire run (see Figure 7.5). We removed the data from 3 participants as their results were below the chance level which represents less than 20% of the population, this observation was consistent with the literature about BCI illiteracy [Allison et al., 2010]. As the percentage of correct responses did not follow a normal distribution we performed a Friedman test approximated by the χ^2 test with 3 degrees of freedom (4 runs -1). The Friedman test showed a significant effect of the display strategy $[\chi^2(3) = 12.23; p < 0.01]$.

After pairwise comparisons using the Wilcoxon signed rank test, on the effect of the condition on the results, we observed a significant difference between the conditions DS0 and DS15 $[p < 0.05]$ and between DS0 and DS22 $[p < 0.05]$ but not significant between DS0 and DS29. In addition, the mean results of the participants were above chance level (45% [Müller-Putz et al., 2008a]) for all of the evaluated display strategies.

The main result of this experiment is that the targets display strategy preferred by participants (DS29) in user study 1, is also performing very well in terms of BCI performance, being similar to the control and optimal display strategy DS0. We hypothesize that, as the only common

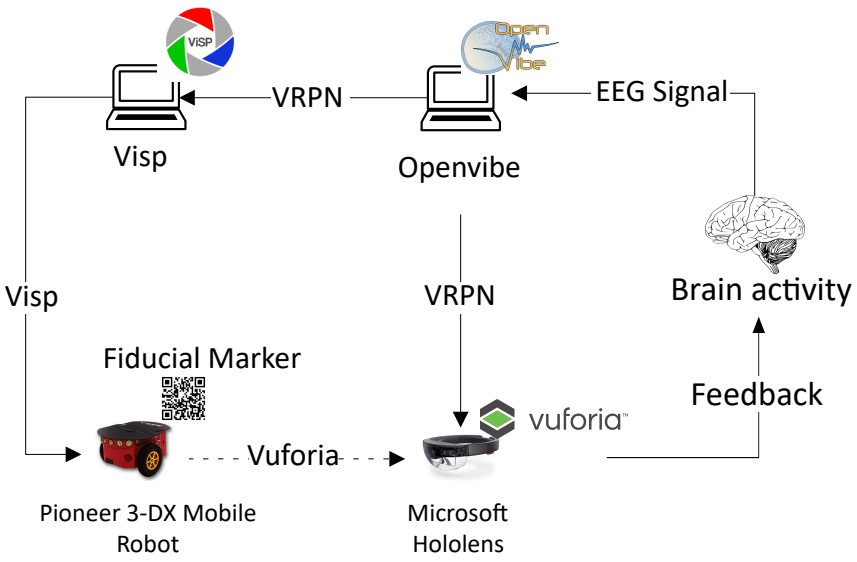


Figure 7.6: Prototype architecture: Brain activity is measured through EEG and analyzed on a computer running the OpenViBE software. The classification results are sent through VRPN to the HoloLens for feedback display and to a computer running Visp software for controlling the robot. The detected class is translated into a Visp command to make the robot move in the desired direction. A fiducial marker is placed on the Pioneer 3-DX robot to contextually make the flickering targets appear on the HoloLens when detected with Vuforia.

dimension between DS0 and DS29 is the size of the targets, the loss of performance in DS15 and DS22 is probably due the smaller-sized targets in these two conditions, which is also consistent with the literature [Ng et al., 2012].

In this experiment, two aspects were tested. First, the analysis of the questionnaires about the participants' preferences confirmed what was found in the previous experiment: condition DS29 was significantly the most preferred one. Secondly, the quantitative evaluation of the SSVEP classification accuracy shows that the most preferred display strategy (DS29) was not significantly worse than the control strategy (DS0). Taken together our results confirm the recommendation to display the targets in the same plan as the robot's (horizontally) and with an exo-centered frame-of-reference, as this is both preferred by the users and gives the same level of performance as more classical display strategies.

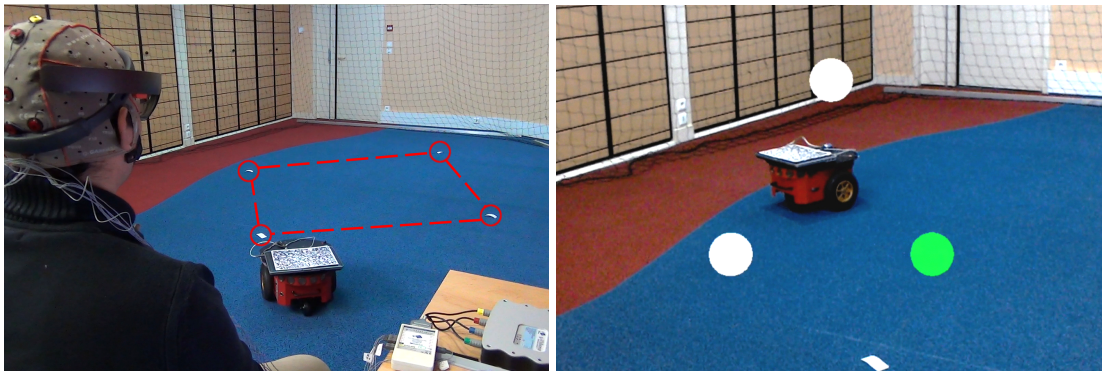


Figure 7.7: Illustration of our final prototype in use. (Left) general overview of the setup with the user equipped with EEG, sitting and facing the real mobile robot. (Right) First person view, as seen from the HoloLens using display strategy DS0. The dashed line represents the path that the robot moved through during testing sessions.

7.5 Final Prototype

The last step of our approach consisted in the development of a final prototype to illustrate and integrate our results in a unique setup. The application being the control of a mobile robot in AR by using a BCI and SSVEP-based commands. We consider here the use of a real mobile robot.

The general architecture of our final prototype is presented in Figure 7.6. Our system includes several hardware elements and software libraries. In addition to the g.tec EEG cap to record brain activity, we used the same Microsoft HoloLens headset. The robot used was a Pioneer 3-DX controlled by an additional computer through an application running a Visp-based [Marchand et al., 2005] application. The hardware setup is presented in Figure 7.7. EEG data was analyzed on-line using OpenViBE which communicated with the Visp [Marchand et al., 2005] application and the AR application running in the HoloLens through VRPN (Virtual-Reality Peripheral Network) [Taylor II et al., 2001] messages. The tracking of the fiducial marker on the robot was done using the Vuforia AR tracking library, compatible with the Universal Windows Platform of the Hololens.

The robot was controlled using the same three directional commands: (1) forward to make it go forward at a speed of 0.2 m/s, (2) rightward to make it do a rightward rotation of 15deg/s and (3) leftward for the leftward rotation at 15 deg/s. The commands consisted in three SSVEP targets. These targets were activated and displayed when the fiducial marker placed on top of the robot was in the field of view of the user. Focusing on one of the targets triggered its

associated command. If the user was not focusing on any target the robot stopped (see Figure 7.7). The control of the prototype robot was continuous. Thus, the user was able to perform head movements while steering the robot. The layout of the targets could be done using any of the display strategies presented in our design space of the previous section. During our testing session, we used DS0.

We conducted informal testings with the real robots. A path was set up with 4 markers on the ground, and participants had to move the robot through each one of them. The targets were displayed following the DS0 display strategy. Informal testings showed that participants were well able to reasonably control the robot along the defined path, given the latency of SSVEP recognition ($\approx 7s$).

7.6 Conclusion

In this Chapter, we proposed an extended design space for target integration in AR-SSVEP based applications. The display strategy relies on five dimensions: orientation, frame-of-reference, anchorage, size and explicitness. A user study conducted with 42 participants enabled to identify and rank the most intuitive/coherent among 32 display strategies. This study notably showed that participants globally preferred when the frame-of-reference of the targets layout was set on the robot (rather than on the user). This could be explained by the spatial rotation task it requires to map targets and directions (represented by the wheels) when they are not in the same frame-of-reference. It also showed that the users preferred when the targets plane was traversal (rather than frontal), i.e. parallel to the 2D robot motion which could be explained by the 2D nature of the task, controlling a mobile robot. In other words, participants could have associated more easily the directions with the targets when the targets were in the same plane as the robot motion. A second study conducted with 24 participants showed that the targets' layout had an effect on the BCI performance. This suggests that the intuitiveness of the display strategy and the task performance have to be balanced.

Finally, this work illustrated the possibility to develop a complete and operational prototype for BCI-based control of a mobile robot in AR. Our final setup makes it possible to steer a real robot in a hands-free manner, opening perspectives for more realistic scenarios involving people with disabilities for instance.

CHAPTER 8

COMBINING BCI AND AR FOR HOME AUTOMATION

Abstract:

This Chapter aims at providing general guidelines towards the design and development of a system combining AR and BCI for Home Automation. It starts by introducing the objective and specifications that such a system should meet. Then it proposes a generic and modular software architecture which can be maintained and adapted for different types of BCI, AR displays and connected objects. Finally, it illustrates how such an architecture can be implemented, through describing the prototype developed and integrated to a commercially available home automation platform.

8.1 Introduction

The fast increasing number of connected objects [Porter and Heppelmann, 2015] and the fact that they are highly heterogeneous, from the connected light, to the smart televisions [Costa et al., 2012] and smart doors [Hentrup et al., 2016], offers a high number of use cases, adapted to the diversity of BCI paradigms. In addition to this, an increased interest can be observed for adaptive houses that are able to sense the inhabitants' states and proactively react to that, as well as for the development of new interaction paradigms for smart homes.

However, from the results of the review conducted in *Chapter 2* about AR and BCIs for smart homes, it appears that most of the so-far developed prototypes were realized for lab-controlled experimentations and never exceeded the state of feasibility studies. Up to now, no developed system was able to cross the borders of the laboratories. Hence the questions: What are the requirements and specifications that an AR-BCI home automation system should meet before we can foresee a broadly available system? How to develop such a system to make it operational and useful to the end-users?

It remains difficult to argue that Brain-Computer Interfaces are ready for large scale exploitation. Their accuracy, speed and affordability do not make BCIs a competitive option with regards to available alternatives. Considering the current state of BCI development, a thorough design and realistic specifications are necessary.

In this Chapter, our goal is to propose general guidelines towards the development of AR-BCI smart-home systems, in a way that they would be operational and possible to upgrade with future technological and scientific improvements. The guidelines presented hereafter, are introduced following a top-down approach. We start by discussing the general, functional and technical specifications: what are the technical requirement and the features the system should provide? Then we present a generic system architecture, describing its different components and their interaction. And finally, we illustrate the implementation of this architecture by describing a specific operational prototype that was developed with the *Orange* company, in an industrial context.

8.2 System Specifications

A smart-home is "*a dwelling incorporating a communication network that connects the key electrical appliances and services, and allows them to be remotely controlled, monitored or accessed*" [Li Jiang et al., 2004]. Like any other smart-home, an AR-BCI system should meet the following specifications:

1. The system has to allow the interaction with multiple objects:

   The first requirement of the system, is that it should enable the operation of multiple devices. The underlying requirement, is that the system has to enable the user to select which object s/he wants to interact with. It is possible to identify two categories of objects: (1) Binary state objects, that can only be switched *ON* or *OFF* (a lamp or a smart plug for e.g) ; and (2) Multiple commands/states objects, that can receive more that 1 command or have more than 2 states (a television for e.g). Some objects may even require to send continuous commands instead of discrete ones. In addition to providing a mechanism to select the intended object, the system has to allow the user to send the desired command.

2. The system has to be self-paced:

   When using an interactive system, the user has to be able to interact at will. This requirement states that the user should be able to stop interacting at will, meaning that the system has to be able to detect idle or rest states and avoid as much false positives as possible.

3. The system has to be modular:

   In order for the system to be maintainable, and to benefit from future technological and scientific advancements, the system has to be modular. In other words, its architecture must comprise independent components interacting, that can be individually upgraded without compromising the functioning of the whole system.

4. The system has to require minimal user training:

   Among the properties that would make a system *usable* on a regular basis, is that it is intuitive to apprehend. However, it is known that BCIs are usually tedious to use [Jeunet et al., 2016a]. As the system is ultimately meant to be operated by naive and non-initiated users, one specification is that it has to require little to no training of the users, about how to modulate their brain activity. This requirement conditions the choice of BCI paradigm and/or processing methods.

5. The system has to contextually adapt its AR display:

   In a general case scenario, the number of controllable appliances can be large. Displaying all the possible commands at every moment on the AR display may lead to confusion. For this matter, the system should adapt its AR interface and the displayed commands, to the objects in the field of view of the user.

6. The system has to be hardware independent:

   Close to the modularity specification (spec. N°3), the system had to be independent from any particular type of hardware. The AR interface has to be possibly displayed on any HMD, phone or tablet screen. Similarly, the system has to be compatible with any, or at least most, of the available EEG systems.

8.3 Generic Architecture

Considering the previous specifications, in particular spec. N° 3 and spec. N° 6, we propose a generic architecture based on 4 components: (1) The BCI ; (2) the AR system ; (3) the Home-Automation platform ; and (4) the Middleware. Each one of these components should be as independent as possible from the others. Each component handles one key aspect of the whole system, and communicates through the middleware (see Figure 8.1). They should not, ideally, be affected by internal changes in any of the other components, although some level of dependency is inevitable.

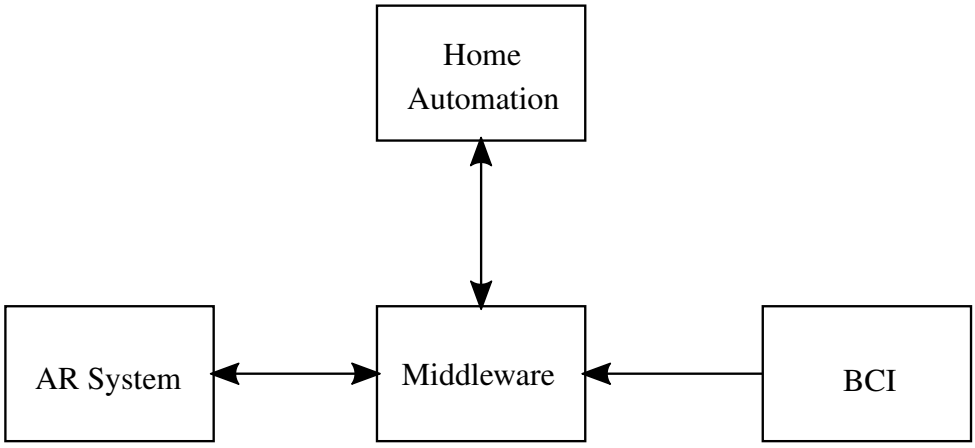


Figure 8.1: Illustration of the components of a generic architecture of a home automation system using BCI and AR. The communication between the elements is done solely through the Middleware.

8.3.1 The BCI

As stated in the definition of a BCI in *Chapter 3*, the role of the BCI is to interpret the brain activity of the user and transmit the corresponding mental state to the interactive system. Depending on the desired BCI paradigm to implement, this component implements the same pipeline described in the Introduction (see Figure 2) in terms of measurement technique (the hardware), signal processing and classification methods. In an operational mode, this component should be able to detect changes in the user's mental state, and send the corresponding information to the middleware.

One reason for decoupling this component is to increase the modularity of the system (spec. N° 3), allowing the designers to change elements (the hardware, signal processing techniques, classification methods) without inducing changes in the rest of the components. As long as the detected mental state is transmitted to the middleware, the whole system should be able to operate.

However, some parts of the component, notably the choice of the BCI paradigm cannot be completely decorrelated from the *AR system* for example. In particular, reactive paradigms rely by definition on the nature and strategy of the stimuli provided by the system.

8.3.2 The AR system

The role of this component is twofold. First, it provides the users with information and feedback about the state of the system, as well as the possible commands to send, depending

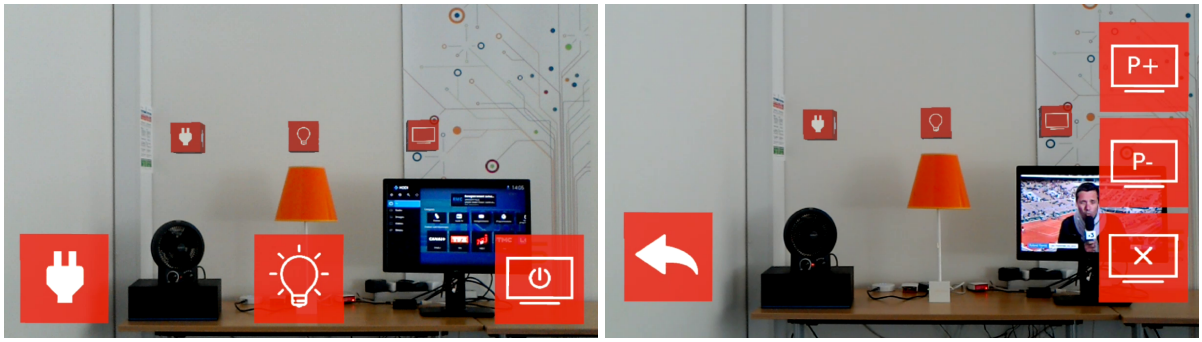


Figure 8.2: Illustration of the implemented AR interface. The default view of the system (Left) represents the different objects in the field of view, with associated flickering icons. The fan and the light could be switched ON or OFF with a single command. The interaction with television was conducted through a hierarchical menu. After selecting the TV, the possible commands to issue appeared on the interface (Right).

on their context (location, objects in field of view for e.g.). As described in *Chapter 7*, where different interfaces and display strategies were evaluated, a particular care should be given to the display strategy so that the system is intuitive and easy to apprehend. Second, in addition to the feedback and information, this component is also responsible of providing the visual stimulation (at least in the case of reactive visual BCI paradigms) for modulating the mental state.

Decoupling this component from the rest of the system allows the designer to change/upgrade the hardware, i.e. the AR display and the AR interface without altering the functioning of the other components.

8.3.3 The Home-Automation platform

This component is the interface between the system and the connected appliances. Its role is to translate the commands obtained from the interpretation of the user mental state determined by the BCI and the context given by the AR system. The nature of this component depends on the specifications and features of the smart objects present in the house, i.e their network protocol, their API [1] and should eventually allow the interaction with heterogeneous objects. If the available smart objects do not share the same protocols, it should be the role of the Home-Automation platform to adapt the commands accordingly.

1. Application Programming Interface

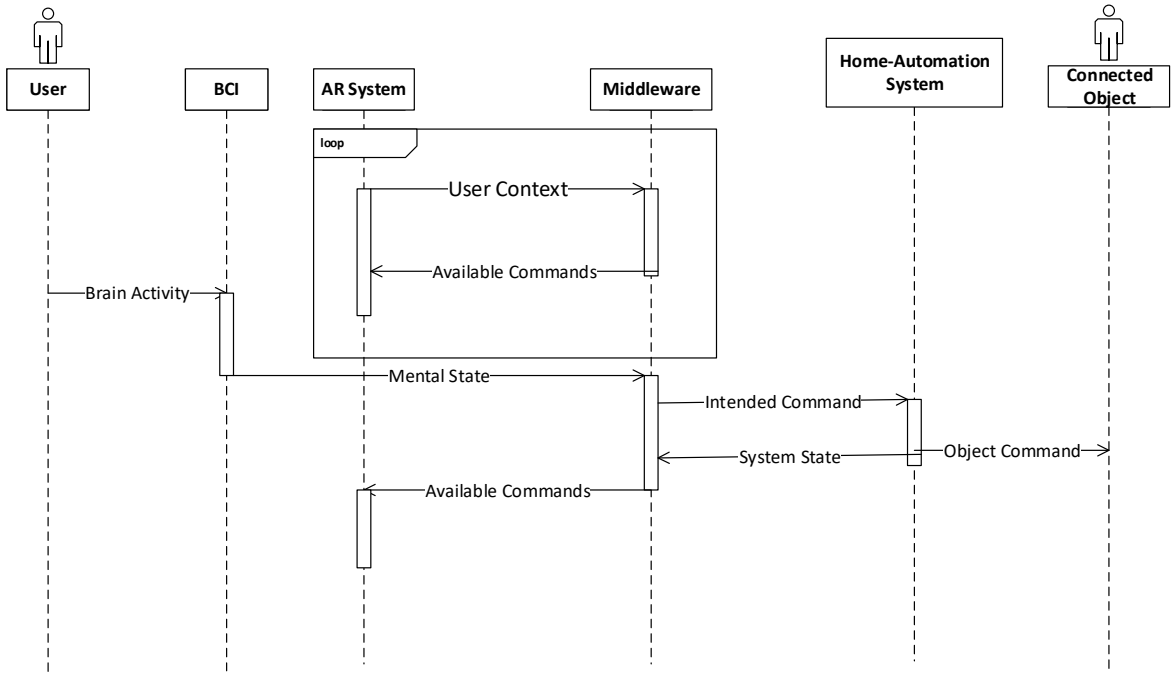


Figure 8.3: Sequence diagram [Garousi et al., 2005] of the system, representing the interactions and the communications between the different system components.

8.3.4 The Middleware

All the components described above have the specificity of being independent from each other. In order to to be able to interact, the system architecture has to provide a communication component. Essentially, the role of the middleware is to determine, from the context of the users and their mental state, the corresponding command to carry to the home automation platform.

The general interaction scenario is described in Figure 8.3. It is the middleware that embeds the "intelligence" of the system as it is responsible for aggregating the information originating from the other components, in order to determine the appropriate command. It is possible to assimilate its functioning to a finite-state automaton. Starting from the initial idle state, the messages coming from the components changes its state and triggers the sending of a command.

Figure 8.4: Illustration of the apparatus used in the developed prototype. (Left) A Microsoft HoloLens and (Right) an mbt Smarting.

8.4 Home Automation Prototype

The aim of this section is to illustrate how the previously described architecture can be implemented for the construction of a functional AR and BCI-based Home Automation system. In particular, we describe the conceptual and technical choices we made in order to meet the requirements mentioned in Section 8.2.

This prototype was built in the scope of an industrial partnership project, with the Orange company, one of the main Internet Service Provider in France. The aim was to build an operational AR-BCI based home automation system, that would be integrated to their commercially available home automation platform *Home'in*.

This project was conducted in collaboration with *Inria Tech*, a service dedicated to the transfer of technology from research centers, to the industry. In this scope, our particular role was to supervise the system design, the BCI development as well as to provide the general guidelines. The development and implementation of the software components, notably the middleware, were conducted by dedicated software engineers.

In addition to the already discussed generic specifications, more specific objectives of the project were added in agreement with the Orange company as follows:

1. The system had to enable the control of at least 3 types of objects.

   In order to be able to showcase the system, the requirement was that the developed system would be able to interact online with 3 objects present in a scene: a fan (connected to a smart plug), a connected light and a Television.

2. The system had to integrate with an existing commercially available home automation system.

   As a collaborative project, the objective was to integrate the global system with the home automation platform provided by Orange.

3. Commands should require at most 2s to be issued.

   The requirement in terms of performance, was that the system had to be able to issue commands in 2s.

4. The system has to be validated online.

   In order to be able to show case the system, and to validate the developments, the demonstration had to be conducted online.

In the following parts, we describe the implementation of each component of the generic architecture, with regards to the previously introduced requirements.

8.4.1 The BCI

Considering the number of objects, and the possible number of commands that the system had to offer (spec N°1), and considering its low requirement in terms of user training (spec N°4) the SSVEP BCI paradigm appeared to be suited for our system.

Therefore, the goal of the BCI component was to determine online, the flickering stimulation frequency that the user was attending and transmit this information to the middleware.

The different possible flickering stimulations and their corresponding commands were stored in a configuration file and managed by the Middleware.

Electroencephalographic (EEG) signals were recorded using a SMARTING amplifier (mBrain-Train, Serbia) with a sampling frequency of 500 Hz and using 5 scalp electrodes: O1, O2, POz, Pz and CPz referenced to Cz and grounded to AFz (see Figure 8.4).

Before being able to use the system online, multi-class CSP and LDA classifiers were calibrated to recognize 5 classes: 4 stimulation frequencies (10Hz, 12Hz, 15Hz and 17Hz) and the idle state, meeting the system's requirement of asynchrony.

The signal acquired online was sliced into 4s long epochs every 2s, meaning that a decision from the classifier was issued every 2s. The power of the signal at the neighbourhood of each stimulation frequency was estimated and combined to form a feature vector. The trained CSPs were applied to spatially filter the data and LDA was used for classification. All implementation was done using OpenViBE [Renard et al., 2010].

Figure 8.5: Illustration of the objects used to showcase the system. The fan (Left) was connected to a smart plug and the connected Light (Middle) and could be turned ON or OFF with a single command. The television (Right) could be controlled using a hierarchical menu.

8.4.2 The AR System

As mentioned previously, and given the employed BCI paradigm (SSVEP), the first objective of this component was to provide the AR interface displaying the flickering SSVEP targets, and to provide the user information about the system state. Its second objective, was to provide the contextual information about the objects detected in the user's field of view (spec N°5).

The AR interface was displayed on a Microsoft Hololens (see Figure 8.4) using the DS0 display strategy (See section 7.3.2) implemented using Unity (see Figure 8.2).

8.4.3 The Home Automation Platform

The commands carried out from the middleware were transmitted to the *Home'In* smart home platform [2]. The selected objects for the prototype were: a Phillips Hue connected lamp and a connected plug on which a fan was plugged, to illustrate the interaction with binary state objects. And a smart television was used to illustrate the interaction with multi-commands objects. The apparatus for the home automation platform is illustrated in Figure 8.5

8.4.4 The Middleware

The middleware was developed using Python. It consisted in several listeners awaiting messages from the BCI component and the AR system, and relaying the appropriate messages to the components.

The implemented architecture of the developed system is summarized in Figure 8.6.

All binary states objects were considered as switches. A single SSVEP selection switches between their 2 states. In the case of multiple commands objects, the interaction was done

2. A commercially available Home Automation platform from Orange, an Internet Service Provider in France.

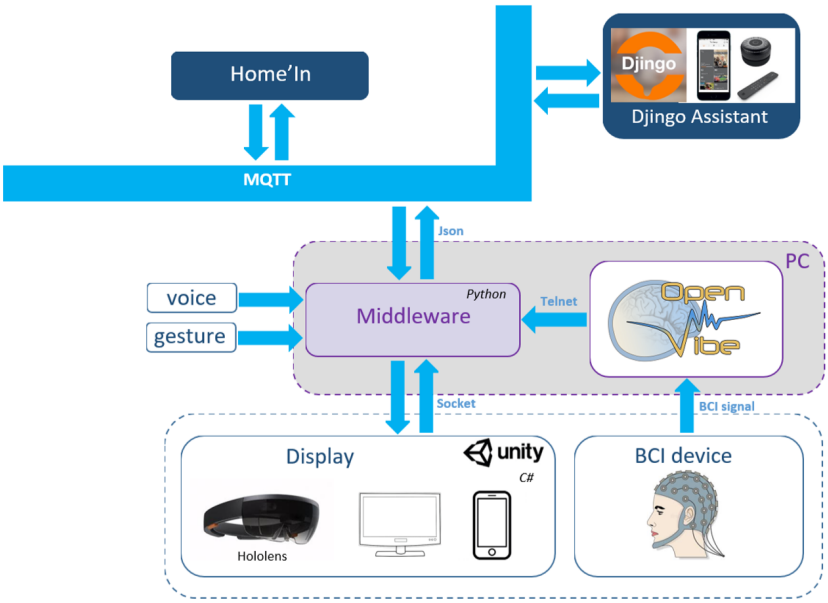


Figure 8.6: Illustration of the final prototype architecture. The components are the same as proposed in the generic architecture, and all communications were done through the middleware. The BCI system was developed using OpenViBE. The middleware was developed in Python, and the Ar system was developed using Unity. The BCI, the AR system and the middleware communicates using TCP [3]sockets.

through a hierarchical menu. First selecting the object, then selecting the available commands (spec N°1).

Depending on the number of available commands at a certain time, and the context provided by the AR display, the middleware associates a flickering frequency to each object's icon and informs the AR system to display the flickering targets. When the BCI notifies the Middleware that a SSVEP command has been recognized, the Middleware notifies the AR system to update the display with the new objects states and change presented icons. The middleware also determines to what object and/or command, the SSVEP stimulation is associated and notifies the smart-home platform.

The middleware allows to meet the specifications 6, 3. Regardless of the EEG system, as long as the BCI provides a detected frequency, the middleware can perform the required action and update the display regardless of its hardware type. In addition, as its role is to receive, interpret then retransmit commands, the middleware can also potentially enable to interact using other modalities (hand gestures, voice etc.) as long as they are simply used to determine the user's intent.

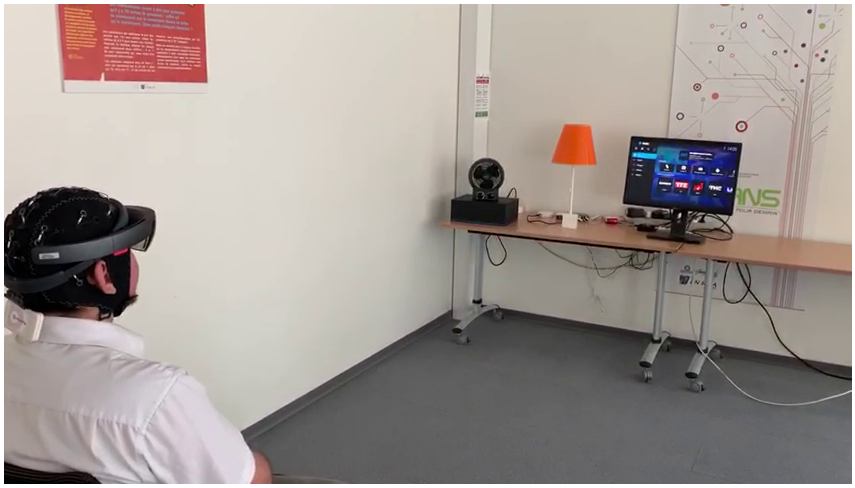


Figure 8.7: Illustration of the final prototype. The user can interact online with the three objects in front of him.

8.5 Informal Testings

The results of our informal testings on 2 participants conducted online (see Figure 8.7), confirmed that the system was functional and allowed to run through a scenario of approximately 10 commands. One participant even achieved the scenario with perfect accuracy.

8.6 Conclusion

In this Chapter we have presented guidelines that support the implementation of the prototype we have designed in order to apply our hypotheses on combining BCI and AR in a real use case. We have chosen Home Automation, as it constitutes one of the fields that offer the most heterogeneous application scenarios, considering the diversity of existing home appliances.

We have provided general specifications that a system combining AR and BCI should meet, and proposed a generic and modular architecture, to facilitate the design of operational prototypes using BCI and AR for Home Automation. These guidelines and architecture were aggregated into an operational system, that was integrated to a commercially available Home Automation platform.

3. Transfer Communication Protocol

CONCLUSION

The global objective of this work was to study the use of BCIs in AR. In particular, the underlying questions in terms of BCI paradigms, of new types of interfaces and system usability. Pursuing this goal, we have identified 5 main objectives:

1. Introducing a new perspective of BCIs.
2. Studying the combination of BCIs and AR in terms of BCI paradigms.
3. Improving the BCI itself.
4. Designing new interfaces for using a BCI in AR.
5. Developing new usages of BCIs and AR.

In order to meet these objectives, we proposed in **Part II**, a new definition of BCIs that highlights the possibility to transform brain activity into inputs of computer processes (in **Chapter 3**).

In **Part III**, we have studied (in **Chapter 4**) the use of two BCI paradigms in AR. We have first assessed the possibility to use the SSVEP paradigm in AR. We have studied the effect of using an OST-HMD on the elicitation and detection of SSVEP responses, showing that displaying an SSVEP interface in AR did not significantly impair the detection accuracy of the SSVEP responses. We have also studied the effect of head movement on the SSVEP paradigms, showing that although high amplitude and speed head movements significantly deteriorated the classification accuracy of the SSVEP responses, low intensity movements were still tolerable. Taken together, there results suggest that it is possible to accurately use the SSVEP paradigm in an AR environment.

Secondly (in **Chapter 5**), we have studied the possibility to elicit and detect ErrPs in AR. In particular, we have shown that tracking errors in Virtual/Augmented Reality do elicit strong ErrPs, and that it was possible to accurately detect the presence of these ErrPs on a single trial basis. This paves the way to new methods for automatically detecting and correcting rendering and tracking issues in Augmented/Virtual Reality.

In **Part IV**, we have deepened the study of the use of the SSVEP paradigm in AR following an approach consisting in: (1) improving the performance of the SSVEP-based BCI itself, (2) improving the intuitiveness of the SSVEP user interface in AR, and (3) improving the application

using an SSVEP-based BCI in AR. In particular, we have proposed (**in Chapter 6**) a novel approach called *HCCA* exploiting the dynamics of Canonical Correlation Coefficients through trained Hidden Markov Models for self-paced detection of SSVEP responses. The results of our testings showed a significant improvement of HCCA over state of the art methods, especially in terms of TNR, which suggests that HCCA better characterizes the "no-control" states of the users while showing comparable accuracy in classifying between the control conditions (SSVEP reponses).

Then (in **Chapter 7**), we have characterized the design space and studied the effect of SSVEP user interfaces in AR, both in terms of intuitiveness and performance. We have chosen the illustrative task of steering a robot on a horizontal plane. We have evaluated the user's perception of the different proposed display strategies and showed that two dimensions of the design space were the most prominent in terms of perceived intuitiveness and coherence, namely *the frame of reference* and *the orientation*. The evaluation of the display strategies corresponding to the highlighted dimensions have shown that the preferred display from the users had comparable performance with standard, BCI-optimal display strategies. These results emphasize the importance of carefully designing the BCI visual interfaces to maximize the trade-off between the BCI performance and the overall usability of the system, showing that it was possible to design intuitive and efficient SSVEP interfaces in AR.

Finally (in **Chapter 8**), we have proposed general guidelines and specifications towards the development of AR BCI-based home automation systems. We have proposed a generic architecture to facilitate the development and implementation of an SSVEP-based AR smart home. Then we have described the prototype that we developed with an industrial partner, implementing the proposed guidelines and architecture, and integrating an AR-BCI system to their commercially available home automation platform.

These results suggest that combining a BCI with an AR system can be used for a variety of purposes. They also illustrate how these two technologies can benefit from each other, with BCI being a new source of information about the user mental state, and AR providing BCIs with an integrated and engaging interface to interact with real and virtual elements.

Our approach provided of insights on how to design BCI-based AR systems, at the same time it also raises numerous challenges to be tackled before such systems can be broadly used.

Future Work and Perspective

The work conducted during this thesis, has left room for improvement as well as some open questions.

As a continuation of this work, we can propose several tracks to explore for improvement. In particular, it would be interesting to study the compatibility of other BCI paradigms with AR environments.

For example, as it was shown that head movements significantly impaired SSVEP classification accuracy (**Chapter 4**), future work could focus on developing more efficient signal filtering methods and artifact removal procedures in order to allow wider and quicker movement while using BCIs. This would for example, enable the selection of quicker moving targets. In addition, in order to increase the usability of the system, it would be interesting to study the possibility of using SSVEP while walking and using an AR-BCI system.

Concerning the detection of ErrPs (**Chapter 5**), the testing and validation were conducted offline, where the beginning and the end of the considered epochs were defined *a priori*. As a future work, we would like to improve our detection method, to asynchronously detect the presence of ErrPs. One way to do this, would be to use sliding windows and continuously classify the recorded EEG epochs. This approach has already been proven effective [Lopes-Dias et al., 2018], and a short-term improvement of our system would be to implement this approach to detect and correct tracking errors in real-time.

The same observation can be made regarding HCCA (**Chapter 6**). To adapt the method to work online, a possible improvement might be to train the temporal classifier to detect transition states, from "no-control" to "selection" and from "selection" to "no-control". The main advantage of this approach would be the potential it offers for continuous SSVEP interaction. If it is possible to detect the moment when the user start sending a command, and keep the active state until a transition to "no-control" state is detected, it would be possible to send a continuous command through the BCI, instead of punctual discrete commands.

We would also extend the proposed design space of SSVEP AR interfaces (**Chapter 7**) to more tasks in addition to mobile robot steering. For example, the question of integrating targets to control 3D moving objects (e.g. flying drones) or multi-function devices would be interesting and important to tackle, in order to make AR-BCI based interaction useful in other application cases.

Lastly, it would also be interesting to compare the performance of the developed AR-BCI home automation system (**Chapter 8**) with other existing alternatives. The whole system would

also benefit from the exploitation of better performing self-paced SSVEP classification methods. As a future work, we intend to include an online version of HCCA for the self-paced detection of the user's commands.

Up to the days where these lines are written, we are very far from any Matrix-like or thought-based interaction machine. Despite clear and significant progress on the analysis and the decoding of brain activity to infer computer commands, seamlessly and intuitively interacting with the brain is still hard. And the question remains: Will we eventually get there? "*Someday these answers may seem as obvious to us as the earth orbiting the sun - or perhaps as ridiculous as a tower of tortoise. Only time (whatever that may be) will tell*" [Hawking, 2009].

BIBLIOGRAPHY

[Acar et al., 2014] Acar, D., Miman, M., and Akirmak, O. O. (2014). Treatment of anxiety disorders patients through eeg and augmented reality. *European Social Sciences Research Journal*, 3(2):18–27.

[Afergan et al., 2015] Afergan, D., Hincks, S. W., Shibata, T., and Jacob, R. J. (2015). Phylter: a system for modulating notifications in wearables using physiological sensing. In *International conference on augmented cognition*, pages 167–177. Springer.

[Allison et al., 2010] Allison, B., Luth, T., Valbuena, D., Teymourian, A., Volosyak, I., and Graser, A. (2010). Bci demographics: How many (and what kinds of) people can use an ssvep bci? *IEEE transactions on neural systems and rehabilitation engineering*, 18(2):107–116.

[Allison et al., 2007] Allison, B. Z., Wolpaw, E. W., and Wolpaw, J. R. (2007). Brain–computer interface systems: progress and prospects. *Expert review of medical devices*, 4(4):463–474.

[Angrisani et al., 2018] Angrisani, L., Arpaia, P., Moccaldi, N., and Esposito, A. (2018). Wearable augmented reality and brain computer interface to improve human-robot interactions in smart industry: A feasibility study for ssvep signals. In *2018 IEEE 4th International Forum on Research and Technology for Society and Industry (RTSI)*, pages 1–5. IEEE.

[Antonenko et al., 2010] Antonenko, P., Paas, F., Grabner, R., and Van Gog, T. (2010). Using electroencephalography to measure cognitive load. *Educational Psychology Review*, 22(4):425–438.

[Azuma et al., 2001] Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., and MacIntyre, B. (2001). Recent advances in augmented reality. *IEEE computer graphics and applications*, 21(6):34–47.

[Azuma, 1997] Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators & Virtual Environments*, 6(4):355–385.

[Balinski and Laraki, 2011] Balinski, M. and Laraki, R. (2011). *Majority judgment: measuring, ranking, and electing*. MIT press.

[baron Fourier, 1822] baron Fourier, J. B. J. (1822). *Théorie analytique de la chaleur*. F. Didot.

[Barresi et al., 2015] Barresi, G., Olivieri, E., Caldwell, D. G., and Mattos, L. S. (2015). Brain-controlled ar feedback design for user's training in surgical hri. In *2015 IEEE International Conference on Systems, Man, and Cybernetics*, pages 1116–1121. IEEE.

[Benaroch et al., 2019] Benaroch, C., Jeunet, C., and LOTTE, F. (2019). Are users' traits informative enough to predict/explain their mental-imagery based BCI performances ? In *8th Graz BCI Conference 2019*, Graz, Austria.

[Berger, 1929] Berger, H. (1929). Über das elektrenkephalogramm des menschen. *European archives of psychiatry and clinical neuroscience*, 87(1):527–570.

[Billinghurst et al., 2015] Billinghurst, M., Clark, A., Lee, G., et al. (2015). A survey of augmented reality. *Foundations and Trends® in Human–Computer Interaction*, 8(2-3):73–272.

[Birbaumer et al., 2003] Birbaumer, N., Hinterberger, T., Kubler, A., and Neumann, N. (2003). The thought-translation device (ttd): neurobehavioral mechanisms and clinical outcome. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2):120–123.

[Blankertz et al., 2006] Blankertz, B., Dornhege, G., Krauledat, M., Muller, K.-R., Kunzmann, V., Losch, F., and Curio, G. (2006). The berlin brain-computer interface: Eeg-based communication without subject training. *IEEE transactions on neural systems and rehabilitation engineering*, 14(2):147–152.

[Blankertz et al., 2007] Blankertz, B., Kawanabe, M., Tomioka, R., Hohlefeld, F. U., Nikulin, V. V., and Müller, K.-R. (2007). Invariant common spatial patterns: Alleviating nonstationarities in brain-computer interfacing. In *NIPS*, pages 113–120.

[Blankertz et al., 2011] Blankertz, B., Lemm, S., Treder, M., Haufe, S., and Müller, K.-R. (2011). Single-trial analysis and classification of erp components — a tutorial. *NeuroImage*, 56(2):814 – 825. Multivariate Decoding and Brain Reading.

[Blum et al., 2012] Blum, T., Stauder, R., Euler, E., and Navab, N. (2012). Superman-like x-ray vision: Towards brain-computer interfaces for medical augmented reality. In *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 271–272.

[Borges et al., 2016] Borges, L. R., Martins, F. R., Naves, E. L., Bastos, T. F., and Lucena, V. F. (2016). Multimodal system for training at distance in a virtual or augmented reality environment for users of electric-powered wheelchairs. *IFAC-PapersOnLine*, 49(30):156–160.

[Broca, 1861] Broca, P. (1861). Remarques sur le siège de la faculté du langage articulé, suivies d'une observation d'aphémie (perte de la parole). *Bulletin et Memoires de la Societe anatomique de Paris*, 6:330–357.

[Cabestaing and Derambure, 2016] Cabestaing, F. and Derambure, P. (2016). *Physiological Markers for Controlling Active and Reactive BCIs*, chapter 4, pages 67–84. John Wiley & Sons, Ltd.

[Cantero et al., 2003] Cantero, J. L., Atienza, M., Stickgold, R., Kahana, M. J., Madsen, J. R., and Kocsis, B. (2003). Sleep-dependent θ oscillations in the human hippocampus and neocortex. *Journal of Neuroscience*, 23(34):10897–10903.

[Caton, 1875] Caton, R. (1875). Electrical currents of the brain. *The Journal of Nervous and Mental Disease*, 2(4):610.

[Cecotti, 2010] Cecotti, H. (2010). A self-paced and calibration-less ssvep-based brain–computer interface speller. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(2):127–133.

[Chapman and Bragdon, 1964] Chapman, R. M. and Bragdon, H. R. (1964). Evoked responses to numerical and non-numerical visual stimuli while problem solving. *Nature*, 203(4950):1155.

[Chavarriaga et al., 2014] Chavarriaga, R., Sobolewski, A., and Millán, J. d. R. (2014). Errare machinale est: the use of error-related potentials in brain-machine interfaces. *Frontiers in neuroscience*, 8:208.

[Chen et al., 2014] Chen, X., Chen, Z., Gao, S., and Gao, X. (2014). A high-itr ssvep-based bci speller. *Brain-Computer Interfaces*, 1(3-4):181–191.

[Chen et al., 2015] Chen, X., Wang, Y., Nakanishi, M., Gao, X., Jung, T.-P., and Gao, S. (2015). High-speed spelling with a noninvasive brain–computer interface. *Proceedings of the national academy of sciences*, 112(44):E6058–E6067.

[Cheng et al., 2002] Cheng, M., Gao, X., Gao, S., and Xu, D. (2002). Design and implementation of a brain-computer interface with high transfer rates. *IEEE transactions on biomedical engineering*, 49(10):1181–1186.

[Chin et al., 2010] Chin, Z. Y., Ang, K. K., Wang, C., and Guan, C. (2010). Online performance evaluation of motor imagery bci with augmented-reality virtual hand feedback. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, pages 3341–3344. IEEE.

[Clerc, 2016] Clerc, M. (2016). *Electroencephalography Data Preprocessing*, chapter 6, pages 101–125. John Wiley & Sons, Ltd.

[Combrisson and Jerbi, 2015] Combrisson, E. and Jerbi, K. (2015). Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical

assessment of decoding accuracy. *Journal of Neuroscience Methods*, 250:126 – 136. Cutting-edge EEG Methods.

[Comon and Jutten, 2010] Comon, P. and Jutten, C. (2010). *Handbook of Blind Source Separation: Independent component analysis and applications*. Academic press.

[Congedo et al., 2017] Congedo, M., Barachant, A., and Bhatia, R. (2017). Riemannian geometry for eeg-based brain-computer interfaces; a primer and a review. *Brain-Computer Interfaces*, 4(3):155–174.

[Cooley and Tukey, 1965] Cooley, J. W. and Tukey, J. W. (1965). An algorithm for the machine calculation of complex fourier series. *Mathematics of computation*, 19(90):297–301.

[Correa-Agudelo et al., 2015] Correa-Agudelo, E., Hernandez, A. M., Ferrin, C., and Gomez, J. D. (2015). Vilimbs: Improving phantom limb treatment through multisensory feedback. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '15, pages 1313–1318. ACM.

[Cortes et al., 2018] Cortes, G., Marchand, E., Brincin, G., and Lécuyer, A. (2018). Mosart: Mobile spatial augmented reality for 3d interaction with tangible objects. *Frontiers in Robotics and AI*, 5:93.

[Costa et al., 2012] Costa, L. C., Maruffa, A., Carvalho, W., and Zuffo, M. K. (2012). A framework design for connected television. In *2012 IEEE International Conference on Consumer Electronics (ICCE)*, pages 590–591. IEEE.

[Csikszentmihalyi, 1975] Csikszentmihalyi, M. (1975). Beyond boredom and anxiety. san francisco. *CA, US: Jossey-Bass*.

[Daly and Wolpaw, 2008] Daly, J. J. and Wolpaw, J. R. (2008). Brain–computer interfaces in neurological rehabilitation. *The Lancet Neurology*, 7(11):1032–1043.

[Davison et al., 2007] Davison, A. J., Reid, I. D., Molton, N. D., and Stasse, O. (2007). Monoslam: Real-time single camera slam. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (6):1052–1067.

[Delorme et al., 2001] Delorme, A., Makeig, S., and Sejnowski, T. (2001). Automatic artifact rejection for eeg data using high-order statistics and independant component analysis.

[Delorme et al., 2007] Delorme, A., Sejnowski, T., and Makeig, S. (2007). Enhanced detection of artifacts in eeg data using higher-order statistics and independent component analysis. *Neuroimage*, 34(4):1443–1449.

[Dias et al., 2017] Dias, M. C. L., Sburlea, A. I., and Müller-Putz, G. (2017). Asynchronous detection of interaction errors. In *cuttingEEG*.

[Dissanayake et al., 2001] Dissanayake, M. G., Newman, P., Clark, S., Durrant-Whyte, H. F., and Csorba, M. (2001). A solution to the simultaneous localization and map building (slam) problem. *IEEE Transactions on robotics and automation*, 17(3):229–241.

[Escolano et al., 2011] Escolano, C., Antelis, J. M., and Minguez, J. (2011). A telepresence mobile robot controlled with a noninvasive brain–computer interface. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 42(3):793–804.

[Evain et al., 2017] Evain, A., Argelaguet, F., Roussel, N., Casiez, G., and Lécuyer, A. (2017). Can i think of something else when using a bci?: Cognitive demand of an ssvep-based bci. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 5120–5125. ACM.

[Falkenstein et al., 1991] Falkenstein, M., Hohnsbein, J., Hoormann, J., and Blanke, L. (1991). Effects of crossmodal divided attention on late erp components. ii. error processing in choice reaction tasks. *Electroencephalography and clinical neurophysiology*, 78(6):447–455.

[Falkenstein et al., 2000] Falkenstein, M., Hoormann, J., Christ, S., and Hohnsbein, J. (2000). Erp components on reaction errors and their functional significance: a tutorial. *Biological psychology*, 51(2-3):87–107.

[Faller et al., 2017] Faller, J., Allison, B. Z., Brunner, C., Scherer, R., Schmalstieg, D., Pfurtscheller, G., and Neuper, C. (2017). A feasibility study on ssvep-based interaction with motivating and immersive virtual and augmented reality. *arXiv preprint arXiv:1701.03981*.

[Faller et al., 2010] Faller, J., Leeb, R., Pfurtscheller, G., and Scherer, R. (2010). Avatar navigation in virtual and augmented reality environments using an ssvep bci icabb-2010. In *Proceedings of the Brain-Computer Interfacing and Virtual Reality Workshop W*, volume 1.

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

[Fatourechi et al., 2007] Fatourechi, M., Bashashati, A., Ward, R. K., and Birch, G. E. (2007). Emg and eog artifacts in brain computer interface systems: A survey. *Clinical neurophysiology*, 118(3):480–494.

[Ferrez and Millán, 2008a] Ferrez, P. W. and Millán, J. d. R. (2008a). Error-related eeg potentials generated during simulated brain–computer interaction. *IEEE transactions on biomedical engineering*, 55(3):923–929.

[Ferrez and Millán, 2008b] Ferrez, P. W. and Millán, J. d. R. (2008b). Simultaneous real-time detection of motor imagery and error-related potentials for improved bci accuracy. Technical report, Proceedings of the 4th Intl. Brain-Computer Interface Workshop and Training Course (Graz).

[Frank et al., 2005] Frank, M. J., Woroch, B. S., and Curran, T. (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron*, 47(4):495–501.

[Frey et al., 2014] Frey, J., Gervais, R., Fleck, S., Lotte, F., and Hachet, M. (2014). Teegi: tangible eeg interface. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*, pages 301–308. ACM.

[Friman et al., 2007] Friman, O., Volosyak, I., and Graser, A. (2007). Multiple channel detection of steady-state visual evoked potentials for brain-computer interfaces. *IEEE transactions on biomedical engineering*, 54(4):742–750.

[Fukunaga, 1990] Fukunaga, R. (1990). *Statistical pattern recognition*. Academic Press.

[Garousi et al., 2005] Garousi, V., Briand, L. C., and Labiche, Y. (2005). Control flow analysis of uml 2.0 sequence diagrams. In *European Conference on Model Driven Architecture-Foundations and Applications*, pages 160–174. Springer.

[Gehring et al., 1993] Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E., and Donchin, E. (1993). A neural system for error detection and compensation. *Psychological science*, 4(6):385–390.

[Gehrke et al., 2019] Gehrke, L., Akman, S., Lopes, P., Chen, A., Singh, A. K., Chen, H.-T., Lin, C.-T., and Gramann, K. (2019). Detecting visuo-haptic mismatches in virtual reality using the prediction error negativity of event-related brain potentials. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 427:1–427:11, New York, NY, USA. ACM.

[Gergondet et al., 2011] Gergondet, P., Druon, S., Kheddar, A., Hintermüller, C., Guger, C., and Slater, M. (2011). Using brain-computer interface to steer a humanoid robot. In *2011 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pages 192–197. IEEE.

[Gevins and Smith, 2000] Gevins, A. and Smith, M. E. (2000). Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style. *Cerebral cortex*, 10(9):829–839.

[Gevins and Smith, 2006] Gevins, A. and Smith, M. E. (2006). Electroencephalography (eeg) in neuroergonomics. *Neuroergonomics: The brain at work*, pages 15–31.

[Goldin et al., 2006] Goldin, D., Smolka, S. A., and Wegner, P. (2006). *Interactive computation*. Springer.

[Graimann et al., 2009] Graimann, B., Allison, B., and Pfurtscheller, G. (2009). Brain–computer interfaces: A gentle introduction. In *Brain-computer interfaces*, pages 1–27. Springer.

[Hajcak et al., 2005] Hajcak, G., Holroyd, C. B., Moser, J. S., and Simons, R. F. (2005). Brain potentials associated with expected and unexpected good and bad outcomes. *Psychophysiology*, 42(2):161–170.

[Hawking, 2009] Hawking, S. (2009). *A brief history of time: from big bang to black holes*. Random House.

[Hentrup et al., 2016] Hentrup, A. A., Lu, D., and Roldan, P. R. (2016). Wireless authentication of smart doors using rfid.

[Herff et al., 2014] Herff, C., Heger, D., Fortmann, O., Hennrich, J., Putze, F., and Schultz, T. (2014). Mental workload during n-back task—quantified in the prefrontal cortex using fnirs. *Frontiers in human neuroscience*, 7:935.

[Herrmann, 2001] Herrmann, C. S. (2001). Human eeg responses to 1–100 hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. *Experimental brain research*, 137(3-4):346–353.

[Hewett et al., 1992] Hewett, T. T., Baecker, R., Card, S., Carey, T., Gasen, J., Mantei, M., Perlman, G., Strong, G., and Verplank, W. (1992). *ACM SIGCHI curricula for human-computer interaction*. ACM.

[Höllerer and Feiner, 2004] Höllerer, T. and Feiner, S. (2004). Mobile augmented reality. *Telegeoinformatics: Location-based computing and services*, 21.

[Horki et al., 2010] Horki, P., Neuper, C., Pfurtscheller, G., and Müller-Putz, G. (2010). Asynchronous steady-state visual evoked potential based bci control of a 2-dof artificial upper limb. *Biomedizinische Technik/Biomedical Engineering*, 55(6):367–374.

[Huebner et al., 2018] Huebner, D., Verhoeven, T., Mueller, K., Kindermans, P., and Tangermann, M. (2018). Unsupervised learning for brain-computer interfaces based on event-related potentials: Review and online comparison [research frontier]. *IEEE Computational Intelligence Magazine*, 13(2):66–77.

[Hwang et al., 2012] Hwang, H.-J., Lim, J.-H., Jung, Y.-J., Choi, H., Lee, S. W., and Im, C.-H. (2012). Development of an ssvep-based bci spelling system adopting a qwerty-style led keyboard. *Journal of neuroscience methods*, 208(1):59–65.

[Irawati et al., 2006] Irawati, S., Green, S., Billinghurst, M., Duenser, A., and Ko, H. (2006). An evaluation of an augmented reality multimodal interface using speech and paddle gestures. In *International Conference on Artificial Reality and Telexistence*, pages 272–283. Springer.

[Jasper, 1958] Jasper, H. H. (1958). The ten-twenty electrode system of the international federation. *Electroencephalogr. Clin. Neurophysiol.*, 10:370–375.

[Jensen et al., 2007] Jensen, O., Kaiser, J., and Lachaux, J.-P. (2007). Human gamma-frequency oscillations associated with attention and memory. *Trends in neurosciences*, 30(7):317–324.

[Jeunet et al., 2017] Jeunet, C., Bideau, B., Argelaguet, F., Chavarriaga, R., Millan, J. D. R., Lecuyer, A., and Kulpa, R. (2017). Investigating neurophysiological correlates of covert attention in soccer goalkeepers. In *The city of Rennes (France) is immensely honored to host the Vth World Congress in Science and Soccer, after Portland (2014) and Copenhagen (2015) and before Melbourne (2019).*, page 227.

[Jeunet et al., 2016a] Jeunet, C., Jahanpour, E., and Lotte, F. (2016a). Why standard brain-computer interface (bci) training protocols should be changed: an experimental study. *Journal of neural engineering*, 13(3):036024.

[Jeunet et al., 2018] Jeunet, C., Lotte, F., Batail, J.-M., Philip, P., and Franchi, J.-A. M. (2018). Using recent bci literature to deepen our understanding of clinical neurofeedback: A short review. *Neuroscience*, 378:225–233.

[Jeunet et al., 2016b] Jeunet, C., N'Kaoua, B., and Lotte, F. (2016b). Advances in user-training for mental-imagery-based bci control: Psychological and cognitive factors and their neural correlates. In *Progress in brain research*, volume 228, pages 3–35. Elsevier.

[Kalunga et al., 2016] Kalunga, E. K., Chevallier, S., Barthélemy, Q., Djouani, K., Monacelli, E., and Hamam, Y. (2016). Online ssvep-based bci using riemannian geometry. *Neurocomputing*, 191:55–68.

[Kandel and Tollet, 2016] Kandel, M. and Tollet, M. (2016). *Anatomy of the Nervous System*, chapter 1, pages 1–24. John Wiley and Sons, Ltd.

[Kansaku et al., 2010] Kansaku, K., Hata, N., and Takano, K. (2010). My thoughts through a robot's eyes: An augmented reality-brain–machine interface. *Neuroscience research*, 66(2):219–222.

[Kemeny and Snell, 1976] Kemeny, J. G. and Snell, J. L. (1976). *Markov Chains*. Springer-Verlag, New York.

[Kothe, 2014] Kothe, C. (2014). Lab streaming layer (lsl). *https://github.com/sccn/labstreaminglayer. Accessed on October*, 26:2015.

[Kozelka and Pedley, 1990] Kozelka, J. W. and Pedley, T. A. (1990). Beta and mu rhythms. *Journal of Clinical Neurophysiology*, 7(2):191–208.

[Kreilinger et al., 2012] Kreilinger, A., Neuper, C., and Müller-Putz, G. (2012). Error potential detection during continuous movement of an artificial arm controlled by brain-computer interface. *Medical & biological engineering & computing*, 50:223–30.

[Kulke et al., 2016] Kulke, L. V., Atkinson, J., and Braddick, O. (2016). Neural differences between covert and overt attention studied using eeg with simultaneous remote eye tracking. *Frontiers in human neuroscience*, 10:592.

[Lampe et al., 2014] Lampe, T., Fiederer, L. D., Voelker, M., Knorr, A., Riedmiller, M., and Ball, T. (2014). A brain-computer interface for high-level remote control of an autonomous, reinforcement-learning-based robotic system for reaching and grasping. In *Proceedings of the 19th international conference on Intelligent User Interfaces*, pages 83–88. ACM.

[Lee and Choi, 2002] Lee, H. and Choi, S. (2002). Pca-based linear dynamical systems for multichannel eeg classification. In *Proceedings of the 9th International Conference on Neural Information Processing, 2002. ICONIP'02.*, volume 2, pages 745–749. IEEE.

[Legeny et al., 2013] Legeny, J., Viciana-Abad, R., and Lecuyer, A. (2013). Toward contextual ssvep-based bci controller: smart activation of stimuli and control weighting. *IEEE transactions on Computational Intelligence and AI in Games*, 5(2):111–116.

[Lenhardt and Ritter, 2010] Lenhardt, A. and Ritter, H. (2010). An augmented-reality based brain-computer interface for robot control. In *International Conference on Neural Information Processing*, pages 58–65. Springer.

[Li Jiang et al., 2004] Li Jiang, Da-You Liu, and Bo Yang (2004). Smart home research. In *Proceedings of 2004 International Conference on Machine Learning and Cybernetics (IEEE Cat. No.04EX826)*, volume 2, pages 659–663 vol.2.

[Lin et al., 2014] Lin, Y.-P., Wang, Y., Wei, C.-S., and Jung, T.-P. (2014). Assessing the quality of steady-state visual-evoked potentials for moving humans using a mobile electroencephalogram headset. *Frontiers in human neuroscience*, 8.

[Lin et al., 2006] Lin, Z., Zhang, C., Wu, W., and Gao, X. (2006). Frequency recognition based on canonical correlation analysis for ssvep-based bcis. *IEEE transactions on biomedical engineering*, 53(12):2610–2614.

[Liu et al., 2014] Liu, Q., Chen, K., Ai, Q., and Xie, S. Q. (2014). recent development of signal processing algorithms for ssvep-based brain computer interfaces. *Journal of Medical and Biological Engineering*, 34(4):299–309.

[Llera et al., 2012] Llera, A., Gómez, V., and Kappen, H. J. (2012). Adaptive classification on brain-computer interfaces using reinforcement signals. *Neural Computation*, 24(11):2900–2923.

[Llera et al., 2011] Llera, A., van Gerven, M. A., Gómez, V., Jensen, O., and Kappen, H. J. (2011). On the use of interaction error potentials for adaptive brain computer interfaces. *Neural Networks*, 24(10):1120–1127.

[Lopes-Dias et al., 2018] Lopes-Dias, C., Sburlea, A. I., and Müller-Putz, G. R. (2018). Masked and unmasked error-related potentials during continuous control and feedback. *Journal of Neural Engineering*, 15(3):036031.

[Lotte et al., 2018] Lotte, F., Bougrain, L., Cichocki, A., Clerc, M., Congedo, M., Rakotomamonjy, A., and Yger, F. (2018). A review of classification algorithms for eeg-based brain–computer interfaces: a 10 year update. *Journal of neural engineering*, 15(3):031005.

[Lotte and Congedo, 2016] Lotte, F. and Congedo, M. (2016). *EEG Feature Extraction*, chapter 7, pages 127–143. John Wiley & Sons, Ltd.

[Lotte et al., 2007] Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F., and Arnaldi, B. (2007). A review of classification algorithms for eeg-based brain–computer interfaces. *Journal of neural engineering*, 4(2):R1.

[Lotte and Guan, 2010] Lotte, F. and Guan, C. (2010). Regularizing common spatial patterns to improve bci designs: unified theory and new algorithms. *IEEE Transactions on biomedical Engineering*, 58(2):355–362.

[Lotte and Jeunet, 2015] Lotte, F. and Jeunet, C. (2015). Towards improved bci based on human learning principles. In *The 3rd International Winter Conference on Brain-Computer Interface*, pages 1–4. IEEE.

[Malik and Amin, 2017] Malik, A. S. and Amin, H. U. (2017). Chapter 1 - designing an eeg experiment. In Malik, A. S. and Amin, H. U., editors, *Designing EEG Experiments for Studying the Brain*, pages 1 – 30. Academic Press.

[Mandel et al., 2009] Mandel, C., Lüth, T., Laue, T., Röfer, T., Gräser, A., and Krieg-Brückner, B. (2009). Navigating a smart wheelchair with a brain-computer interface interpreting steady-state visual evoked potentials. In *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*, pages 1118–1125. IEEE.

[Maquet et al., 1997] Maquet, P., Degueldre, C., Delfiore, G., Aerts, J., Péters, J.-M., Luxen, A., and Franck, G. (1997). Functional neuroanatomy of human slow wave sleep. *Journal of Neuroscience*, 17(8):2807–2812.

[Marchand et al., 2005] Marchand, É., Spindler, F., and Chaumette, F. (2005). Visp for visual servoing: a generic software platform with a wide class of robot control skills. *IEEE Robotics & Automation Magazine*, 12(4):40–52.

[Martens et al., 2012] Martens, N., Jenke, R., Abu-Alqumsan, M., Kapeller, C., Hintermüller, C., Guger, C., Peer, A., and Buss, M. (2012). Towards robotic re-embodiment using a brain-and-body-computer interface. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5131–5132. IEEE.

[Mason and Birch, 2003] Mason, S. G. and Birch, G. E. (2003). A general framework for brain-computer interface design. *IEEE transactions on neural systems and rehabilitation engineering*, 11(1):70–85.

[McFarland, 2015] McFarland, D. J. (2015). The advantages of the surface laplacian in brain–computer interface research. *International Journal of Psychophysiology*, 97(3):271–276.

[McFarland et al., 1997] McFarland, D. J., McCane, L. M., David, S. V., and Wolpaw, J. R. (1997). Spatial filter selection for eeg-based communication. *Electroencephalography and clinical Neurophysiology*, 103(3):386–394.

[McFarland et al., 2005] McFarland, D. J., Sarnacki, W. A., Vaughan, T. M., and Wolpaw, J. R. (2005). Brain-computer interface (bci) operation: signal and noise during early training sessions. *Clinical Neurophysiology*, 116(1):56–62.

[Mercier-Ganady et al., 2014] Mercier-Ganady, J., Lotte, F., Loup-Escande, E., Marchal, M., and Lécuyer, A. (2014). The mind-mirror: See your brain in action in your head using eeg and augmented reality. In *2014 IEEE Virtual Reality (VR)*, pages 33–38. IEEE.

[Mika et al., 1999] Mika, S., Ratsch, G., Weston, J., Scholkopf, B., and Mullers, K.-R. (1999). Fisher discriminant analysis with kernels. In *Neural networks for signal processing IX: Proceedings of the 1999 IEEE signal processing society workshop (cat. no. 98th8468)*, pages 41–48. Ieee.

[Milekovic et al., 2012] Milekovic, T., Ball, T., Schulze-Bonhage, A., Aertsen, A., and Mehring, C. (2012). Error-related electrocorticographic activity in humans during continuous movements. *Journal of neural engineering*, 9(2):026007.

[Milgram and Kishino, 1994] Milgram, P. and Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, 77(12):1321–1329.

[Missonnier et al., 2006] Missonnier, P., Deiber, M.-P., Gold, G., Millet, P., Pun, M. G.-F., Fazio-Costa, L., Giannakopoulos, P., and Ibáñez, V. (2006). Frontal theta event-related

synchronization: comparison of directed attention and working memory load effects. *Journal of Neural Transmission*, 113(10):1477–1486.

[Motamedi-Fakhr et al., 2014] Motamedi-Fakhr, S., Moshrefi-Torbati, M., Hill, M., Hill, C. M., and White, P. R. (2014). Signal processing techniques applied to human sleep eeg signals—a review. *Biomedical Signal Processing and Control*, 10:21–33.

[Müller et al., 2008] Müller, K.-R., Tangermann, M., Dornhege, G., Krauledat, M., Curio, G., and Blankertz, B. (2008). Machine learning for real-time single-trial eeg-analysis: from brain–computer interfacing to mental state monitoring. *Journal of neuroscience methods*, 167(1):82–90.

[Müller-Putz et al., 2008a] Müller-Putz, G., Scherer, R., Brunner, C., Leeb, R., and Pfurtscheller, G. (2008a). Better than random: a closer look on bci results. *International Journal of Bioelectromagnetism*, 10(EPFL-ARTICLE-164768):52–55.

[Müller-Putz et al., 2005] Müller-Putz, G. R., Scherer, R., Brauneis, C., and Pfurtscheller, G. (2005). Steady-state visual evoked potential (ssvep)-based communication: impact of harmonic frequency components. *Journal of neural engineering*, 2(4):123.

[Müller-Putz et al., 2008b] Müller-Putz, G. R., Scherer, R., Brunner, C., Leeb, R., and Pfurtscheller, G. (2008b). Better than random: a closer look on bci results. *International Journal of Bioelectromagnetism*.

[Nan et al., 2011] Nan, W., Wong, C. M., Wang, B., Wan, F., Mak, P. U., Mak, P. I., and Vai, M. I. (2011). A comparison of minimum energy combination and canonical correlation analysis for ssvep detection. In *Neural Engineering (NER), 2011 5th International IEEE/EMBS Conference on*, pages 469–472. IEEE.

[Navarro-Cebrian et al., 2013] Navarro-Cebrian, A., Knight, R. T., and Kayser, A. S. (2013). Error-monitoring and post-error compensations: dissociation between perceptual failures and motor errors with and without awareness. *Journal of Neuroscience*, 33(30):12375–12383.

[Ng et al., 2012] Ng, K. B., Bradley, A. P., and Cunnington, R. (2012). Stimulus specificity of a steady-state visual-evoked potential-based brain–computer interface. *Journal of Neural engineering*, 9(3):036008.

[Nicolas-Alonso and Gomez-Gil, 2012] Nicolas-Alonso, L. F. and Gomez-Gil, J. (2012). Brain computer interfaces, a review. *sensors*, 12(2):1211–1279.

[Niedermeyer, 2005] Niedermeyer, E. (2005). Electroencephalography. *Basic principles, clinical applications, and related fields*.

[Ogawa et al., 1990] Ogawa, S., Lee, T.-M., Nayak, A. S., and Glynn, P. (1990). Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *Magnetic resonance in medicine*, 14(1):68–78.

[Omedes et al., 2015] Omedes, J., Iturrate, I., Minguez, J., and Montesano, L. (2015). Analysis and asynchronous detection of gradually unfolding errors during monitoring tasks. *Journal of Neural Engineering*, 12(5):056001.

[Omedes et al., 2014] Omedes, J., Iturrate, I., and Montesano, L. (2014). Brain connectivity in continuous error tasks. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 3997–4000.

[Padrao et al., 2016] Padrao, G., Gonzalez-Franco, M., Sanchez-Vives, M. V., Slater, M., and Rodriguez-Fornells, A. (2016). Violating body movement semantics: neural signatures of self-generated and external-generated errors. *Neuroimage*, 124:147–156.

[Pastor et al., 2003] Pastor, M. A., Artieda, J., Arbizu, J., Valencia, M., and Masdeu, J. C. (2003). Human cerebral activation during steady-state visual-evoked responses. *Journal of neuroscience*, 23(37):11621–11627.

[Pedregosa et al., 2011] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

[Perronnet et al., 2017] Perronnet, L., Lécuyer, A., Mano, M., Bannier, E., Lotte, F., Clerc, M., and Barillot, C. (2017). Unimodal versus bimodal eeg-fmri neurofeedback of a motor imagery task. *Frontiers in human neuroscience*, 11:193.

[Petit et al., 2014] Petit, D., Gergondet, P., Cherubini, A., Meilland, M., Comport, A. I., and Kheddar, A. (2014). Navigation assistance for a bci-controlled humanoid robot. In *Cyber Technology in Automation, Control, and Intelligent Systems (CYBER), 2014 IEEE 4th Annual International Conference on*, pages 246–251. IEEE.

[Pezzetta et al., 2018] Pezzetta, R., Nicolardi, V., Tidoni, E., and Aglioti, S. M. (2018). Error, rather than its probability, elicits specific electrocortical signatures: a combined eeg-immersive virtual reality study of action observation. *Journal of neurophysiology*.

[Pfurtscheller et al., 2006] Pfurtscheller, G., Graimann, B., and Neuper, C. (2006). Eeg-based brain-computer interface system. *Wiley Encyclopedia of Biomedical Engineering*.

[Pfurtscheller and Neuper, 2001] Pfurtscheller, G. and Neuper, C. (2001). Motor imagery and direct brain-computer communication. *Proceedings of the IEEE*, 89(7):1123–1134.

[Pfurtscheller et al., 2000] Pfurtscheller, G., Neuper, C., Guger, C., Harkam, W., Ramoser, H., Schlogl, A., Obermaier, B., and Pregenzer, M. (2000). Current trends in graz brain-computer interface (bci) research. *IEEE transactions on rehabilitation engineering*, 8(2):216–219.

[Polich and Margala, 1997] Polich, J. and Margala, C. (1997). P300 and probability: comparison of oddball and single-stimulus paradigms. *International Journal of Psychophysiology*, 25(2):169–176.

[Porter and Heppelmann, 2015] Porter, M. E. and Heppelmann, J. E. (2015). How smart, connected products are transforming companies. *Harvard business review*, 93(10):96–114.

[Poryzala and Materka, 2014] Poryzala, P. and Materka, A. (2014). Cluster analysis of cca coefficients for robust detection of the asynchronous ssveps in brain–computer interfaces. *Biomedical Signal Processing and Control*, 10:201–208.

[Potts et al., 2019] Potts, D., Loveys, K., Ha, H., Huang, S., Billinghurst, M., and Broadbent, E. (2019). Zeng: Ar neurofeedback for meditative mixed reality. In *Proceedings of the 2019 on Creativity and Cognition*, pages 583–590. ACM.

[Purves et al., 2004] Purves, D., Augustine, G., Fitzpatrick, D., Hall, W., Lamantia, A., Mcnamara, J., and Williams, S. (2004). Studying the nervous systems of humans and other animals. *Polirradiculoneurite aguda em cães: estudo descritivo de*, 20.

[Putze et al., 2019] Putze, F., Weiß, D., Vortmann, L.-M., and Schultz, T. (2019). Augmented reality interface for smart home control using ssvep-bci and eye gaze.

[Rabiner, 1989] Rabiner, L. R. (1989). A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286.

[Ramoser et al., 2000] Ramoser, H., Muller-Gerking, J., and Pfurtscheller, G. (2000). Optimal spatial filtering of single trial eeg during imagined hand movement. *IEEE transactions on rehabilitation engineering*, 8(4):441–446.

[Regan, 1977] Regan, D. (1977). Steady-state evoked potentials. *J. Opt. Soc. Am.*, 67(11):1475–1489.

[Regan, 1989] Regan, D. (1989). Human brain electrophysiology. *Evoked potentials and evoked magnetic fields in science and medicine*.

[Renard et al., 2010] Renard, Y., Lotte, F., Gibert, G., Congedo, M., Maby, E., Delannoy, V., Bertrand, O., and Lécuyer, A. (2010). Openvibe: An open-source software platform to design, test, and use brain–computer interfaces in real and virtual environments. *Presence: teleoperators and virtual environments*, 19(1):35–53.

[Rivet et al., 2009] Rivet, B., Souloumiac, A., Attina, V., and Gibert, G. (2009). xdawn algorithm to enhance evoked potentials: application to brain–computer interface. *IEEE Transactions on Biomedical Engineering*, 56(8):2035–2043.

[Roy and Frey, 2016] Roy, R. N. and Frey, J. (2016). *Neurophysiological Markers for Passive Brain–Computer Interfaces*, chapter 5, pages 85–100. John Wiley & Sons, Ltd.

[Salazar-Gomez et al., 2017] Salazar-Gomez, A. F., DelPreto, J., Gil, S., Guenther, F. H., and Rus, D. (2017). Correcting robot mistakes in real time using eeg signals. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6570–6577.

[Schalk et al., 2000] Schalk, G., Wolpaw, J. R., McFarland, D. J., and Pfurtscheller, G. (2000). Eeg-based communication: presence of an error potential. *Clinical Neurophysiology*, 111(12):2138 – 2144.

[Schall et al., 2009] Schall, G., Mendez, E., Kruijff, E., Veas, E., Junghanns, S., Reitinger, B., and Schmalstieg, D. (2009). Handheld augmented reality for underground infrastructure visualization. *Personal and ubiquitous computing*, 13(4):281–291.

[Sherman and Guillery, 2001] Sherman, S. M. and Guillery, R. (2001). Chapter i - introduction. In Sherman, S. M. and Guillery, R., editors, *Exploring the Thalamus*, pages 1 – 18. Academic Press, San Diego.

[Shibata et al., 2014] Shibata, T., Peck, E. M., Afergan, D., Hincks, S. W., Yuksel, B. F., and Jacob, R. J. (2014). Building implicit interfaces for wearable computers with physiological inputs: zero shutter camera and phylter. In *Proceedings of the adjunct publication of the 27th annual ACM symposium on User interface software and technology*, pages 89–90. ACM.

[Si-Mohammed et al., 2018] Si-Mohammed, H., Petit, J., Jeunet, C., Argelaguet, F., Spindler, F., Évain, A., Roussel, N., Casiez, G., and Lécuyer, A. (2018). Towards bci-based interfaces for augmented reality: Feasibility, design and evaluation. *IEEE transactions on visualization and computer graphics*.

[Spüler and Niethammer, 2015a] Spüler, M. and Niethammer, C. (2015a). Error-related potentials during continuous feedback: using eeg to detect errors of different type and severity. *Frontiers in Human Neuroscience*, 9:155.

[Spüler and Niethammer, 2015b] Spüler, M. and Niethammer, C. (2015b). Error-related potentials during continuous feedback: using eeg to detect errors of different type and severity. *Frontiers in Human Neuroscience*, 9:155.

[Strehl et al., 2006] Strehl, U., Leins, U., Goth, G., Klinger, C., Hinterberger, T., and Birbaumer, N. (2006). Self-regulation of slow cortical potentials: A new treatment for children with attention-deficit/hyperactivity disorder. *Pediatrics*, 118(5):e1530–e1540.

[Suefusa and Tanaka, 2018] Suefusa, K. and Tanaka, T. (2018). Asynchronous brain–computer interfacing based on mixed-coded visual stimuli. *IEEE Transactions on Biomedical Engineering*, 65(9):2119–2129.

[Sugiarto et al., 2009] Sugiarto, I., Allison, B., and Gräser, A. (2009). Optimization strategy for ssvep-based bci in spelling program application. In *2009 International Conference on Computer Engineering and Technology*, volume 1, pages 223–226. IEEE.

[Takano et al., 2011] Takano, K., Hata, N., and Kansaku, K. (2011). Towards intelligent environments: An augmented realityâ€“brainâ€“machine interface operated with a see-through head-mount display. *Frontiers in Neuroscience*, 5:60.

[Taylor II et al., 2001] Taylor II, R. M., Hudson, T. C., Seeger, A., Weber, H., Juliano, J., and Helser, A. T. (2001). Vrpn: a device-independent, network-transparent vr peripheral system. In *Proceedings of the ACM symposium on Virtual reality software and technology*, pages 55–61. ACM.

[Thrun and Leonard, 2008] Thrun, S. and Leonard, J. J. (2008). Simultaneous localization and mapping. *Springer handbook of robotics*, pages 871–889.

[Tonin et al., 2013] Tonin, L., Leeb, R., Sobolewski, A., and del R Millán, J. (2013). An online eeg bci based on covert visuospatial attention in absence of exogenous stimulation. *Journal of neural engineering*, 10(5):056007.

[Tonin et al., 2010] Tonin, L., Leeb, R., Tavella, M., Perdikis, S., and Millán, J. d. R. (2010). The role of shared-control in bci-based telepresence. In *2010 IEEE International Conference on Systems, Man and Cybernetics*, pages 1462–1466. IEEE.

[Trunk, 1979] Trunk, G. V. (1979). A problem of dimensionality: A simple example. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-1(3):306–307.

[Uno et al., 2015] Uno, K., Naito, G., Tobisa, Y., Yoshida, L., Ogawa, Y., Kotani, K., and Jimbo, Y. (2015). Basic investigation of brain–computer interface combined with augmented reality and development of an improvement method using the nontarget object. *Electronics and Communications in Japan*, 98(8):9–15.

[van Schie et al., 2004] van Schie, H. T., Mars, R. B., Coles, M. G., and Bekkering, H. (2004). Modulation of activity in medial frontal and motor cortices during error observation. *Nature neuroscience*, 7(5):549.

[Vanderwolf, 2000] Vanderwolf, C. (2000). Are neocortical gamma waves related to consciousness? *Brain research*, 855(2):217–224.

[Vaughan et al., 2006] Vaughan, T. M., McFarland, D. J., Schalk, G., Sarnacki, W. A., Krusienski, D. J., Sellers, E. W., and Wolpaw, J. R. (2006). The wadsworth bci research and development program: at home with bci. *IEEE transactions on neural systems and rehabilitation engineering*, 14(2):229–233.

[Vialatte et al., 2010] Vialatte, F.-B., Maurice, M., Dauwels, J., and Cichocki, A. (2010). Steady-state visually evoked potentials: focus on essential paradigms and future perspectives. *Progress in neurobiology*, 90(4):418–438.

[Vidal, 2016] Vidal, F. (2016). *Cerebral Electrogenesis*, chapter 3, pages 45–65. John Wiley and Sons, Ltd.

[Vidal, 1973] Vidal, J. J. (1973). Toward direct brain-computer communication. *Annual review of Biophysics and Bioengineering*, 2(1):157–180.

[Volosyak et al., 2010] Volosyak, I., Valbuena, D., Luth, T., and Gräser, A. (2010). Towards an ssvep based bci with high itr. *IEEE Trans. Biomed. Eng*.

[Wang et al., 2018] Wang, M., Li, R., Zhang, R., Li, G., and Zhang, D. (2018). A wearable ssvep-based bci system for quadcopter control using head-mounted device. *IEEE Access*, 6:26789–26798.

[Wang et al., 2010] Wang, N., Qian, T., Zhuo, Q., and Gao, X. (2010). Discrimination between idle and work states in bci based on ssvep. In *2010 2nd International Conference on Advanced Computer Control*, volume 4, pages 355–358. IEEE.

[Wang et al., 2008] Wang, Y., Gao, X., Hong, B., Jia, C., and Gao, S. (2008). Brain-computer interfaces based on visual evoked potentials. *IEEE Engineering in Medicine and Biology Magazine*, 27(5):64–71.

[Weiskopf et al., 2004] Weiskopf, N., Mathiak, K., Bock, S. W., Scharnowski, F., Veit, R., Grodd, W., Goebel, R., and Birbaumer, N. (2004). Principles of a brain-computer interface (bci) based on real-time functional magnetic resonance imaging (fmri). *IEEE transactions on biomedical engineering*, 51(6):966–970.

[Wessel, 2012] Wessel, J. R. (2012). Error awareness and the error-related negativity: evaluating the first decade of evidence. *Frontiers in Human Neuroscience*, 6:88.

[Winograd, 1978] Winograd, S. (1978). On computing the discrete fourier transform. *Mathematics of computation*, 32(141):175–199.

[Wolpaw, 2007] Wolpaw, J. R. (2007). Brain–computer interfaces as new brain output pathways. *The Journal of physiology*, 579(3):613–619.

[Wolpaw et al., 2002] Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., and Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clinical neurophysiology*, 113(6):767–791.

[Wolpaw et al., 2006] Wolpaw, J. R., Loeb, G. E., Allison, B. Z., Donchin, E., do Nascimento, O. F., Heetderks, W. J., Nijboer, F., Shain, W. G., and Turner, J. N. (2006). Bci meeting 2005-workshop on signals and recording methods. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14(2):138–141.

[Wolpaw and McFarland, 2004] Wolpaw, J. R. and McFarland, D. J. (2004). Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans. *Proceedings of the national academy of sciences*, 101(51):17849–17854.

[Wolpaw et al., 1991] Wolpaw, J. R., McFarland, D. J., Neat, G. W., and Forneris, C. A. (1991). An eeg-based brain-computer interface for cursor control. *Electroencephalography and clinical neurophysiology*, 78(3):252–259.

[Wolpaw and Wolpaw, 2012] Wolpaw, J. R. and Wolpaw, E. W. (2012). Brain-computer interfaces: something new under the sun. *Brain-computer interfaces: principles and practice*, pages 3–12.

[Xia et al., 2013] Xia, B., Li, X., Xie, H., Yang, W., Li, J., and He, L. (2013). Asynchronous brain–computer interface based on steady-state visual-evoked potential. *Cognitive Computation*, 5(2):243–251.

[y Cajal, 1954] y Cajal, S. R. (1954). *Neuron Theory Or Reticular Theory?: Objetive Evidence of the Anatomical Unity of Nerve Cells*. Editorial CSIC-CSIC Press.

[Yazmir and Reiner, 2017] Yazmir, B. and Reiner, M. (2017). I act, therefore i err: Eeg correlates of success and failure in a virtual throwing game. *International Journal of Psychophysiology*, 122:32–41.

[Yazmir et al., 2016] Yazmir, B., Reiner, M., Pratt, H., and Zacksenhouse, M. (2016). Brain responses to errors during 3d motion in a hapto-visual vr. In *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*, pages 120–130. Springer.

[Yuan and He, 2014] Yuan, H. and He, B. (2014). Brain–computer interfaces using sensorimotor rhythms: current state and future perspectives. *IEEE Transactions on Biomedical Engineering*, 61(5):1425–1435.

[Yuksel et al., 2016] Yuksel, B. F., Oleson, K. B., Harrison, L., Peck, E. M., Afergan, D., Chang, R., and Jacob, R. J. (2016). Learn piano with bach: An adaptive learning interface that adjusts task difficulty based on brain state. In *Proceedings of the 2016 chi conference on human factors in computing systems*, pages 5372–5384. ACM.

[Zander and Kothe, 2011] Zander, T. O. and Kothe, C. (2011). Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general. *Journal of neural engineering*, 8(2):025005.

[Zeng et al., 2017] Zeng, H., Wang, Y., Wu, C., Song, A., Liu, J., Ji, P., Xu, B., Zhu, L., Li, H., and Wen, P. (2017). Closed-loop hybrid gaze brain-machine interface based robotic arm control with augmented reality feedback. *Frontiers in Neurorobotics*, 11:60.

[Zhang et al., 2015] Zhang, H., Chavarriaga, R., Khaliliardali, Z., Gheorghe, L., Iturrate, I., and d R Millán, J. (2015). EEG-based decoding of error-related brain activity in a real-world driving task. *Journal of Neural Engineering*, 12(6):066028.

[Zhou et al., 2008] Zhou, F., Duh, H. B.-L., and Billinghurst, M. (2008). Trends in augmented reality tracking, interaction and display: A review of ten years of ismar. In *Proceedings of the 7th IEEE/ACM international symposium on mixed and augmented reality*, pages 193–202. IEEE Computer Society.

[Zhu et al., 2010] Zhu, D., Bieger, J., Molina, G. G., and Aarts, R. M. (2010). A survey of stimulation methods used in ssvep-based bcis. *Computational intelligence and neuroscience*, 2010:1.

LIST OF FIGURES

LIST OF TABLES

AUTHOR'S LIST OF PUBLICATIONS

Journal

– **Hakim Si-Mohammed**, Jimmy Petit, Camille Jeunet, Ferran Argelaguet, Fabien Spindler, Andéol Evain, Nicolas Roussel, Géry Casiez and Anatole Lécuyer. Towards BCI-based Interfaces for Augmented Reality: Feasibility, Design and Evaluation. IEEE Transactions on Visualization and Computer Graphics. 2018.

Conference

– **Hakim Si-Mohammed**, Ferran Argelaguet, Géry Casiez, Nicolas Roussel, Anatole Lécuyer. Defining Brain-Computer Interfaces: A Human-Computer Interaction Perspective. Graz Brain-Computer Interface Conference, 2019.

– **Hakim Si-Mohammed**, , Ferran Argelaguet, Géry Casiez, Nicolas Roussel, Anatole Lécuyer. Brain Computer Interfaces and Augmented Reality: A State of the Art. Graz Brain-Computer Interface Conference, Sep 2017, Graz, Austria.

Submitted

– **Hakim Si-Mohammed**, Catarina Lopes-Dias, Maria Duarte, Camille Jeunet, Ferran Argelaguet, Géry Casiez, Gernot Müller-Putz, Anatole Lécuyer and Reinhold Scherer. Detecting System Errors in Virtual Reality Using EEG through Error Related Potentials. IEEE VR 2020.

In Preparation

– **Hakim Si-Mohammed**, Ferran Argelaguet, Camille Jeunet, Géry Casiez, Anatole Lécuyer and Fabien Lotte. Using Hidden Markov Models and CCA Coefficients Dynamics for Improving Self-Paced SSVEP-based BCI Accuracy. IEEE Transactions on Biomedical Engineering.

Résumé long en Français de la Thèse

Conception et Etude de Systèmes Interactifs basés sur les Interfaces Cerveau-Ordinateur et la Réalité Augmentée

---

INTRODUCTION

Interagir par la force de son esprit est une idée très répandue dans la culture populaire. La communication par la "pensée" ou ou la manipulation d'environnements virtuels par seule activité cérébrale ont été décrites dans de plusieurs réalisations cinématographiques et de nombreuses oeuvres de fiction. (Figure 8.8).

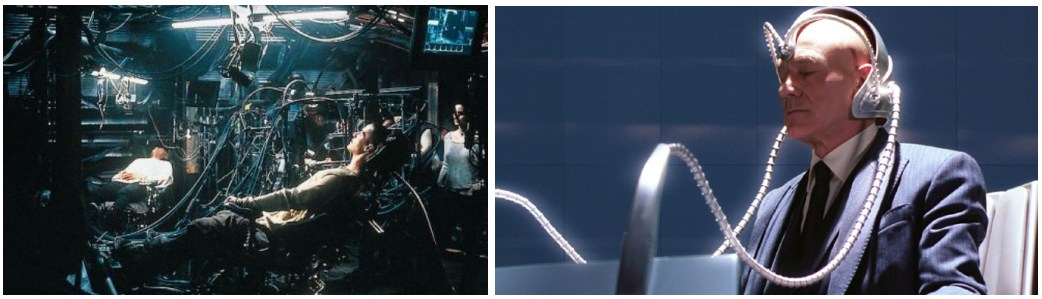


Figure 8.8: Exemples de films employant le concept d'interaction cerveau-ordinateur. Dans le film "Matrix" (à gauche), les protagonistes utilisent une Interface Cerveau-Ordinateur (ICO) invasive pour s'immerger dans un environnement virtuel. Dans le film "X-men" (à droite), l'un des personnages principaux utilise une INCO non-invasive afin de surveilleir et ressentir le monde extérieur.

L'objectif global de ces travaux est d'étudier l'utilisation d'Interfaces Cerveau-Ordinateur (ICO) en Réalité Augmentée (RA). En particulier en termes de paradigmes d'ICO [13], de nouvelles interfaces utilisateurs et d'utilisabilité du système. Les ICO promettent d'interfacer le cerveau et l'ordinateur. La RA quant à elle, vise à intégrer des éléments virtuels dans l'environnement du monde réel [Azuma, 1997]. Ensemble, les ICO et la RA permettent une interaction mains libres avec des éléments virtuels dans le monde réel. Cette interaction peut être utilisée dans de nombreux domaines d'application tels que la robotique, la santé ou le divertissement. Cependant, la combinaison de ces deux domaines scientifiques, ouvre également de nombreuses questions de recherche. Parmi celles-ci : Quel type d'activité cérébrale peut être utilisé dans la RA ? Comment pouvons-nous exploiter les sorties de'une ICO pour interagir avec des éléments réels et virtuels ? Comment construire des systèmes utilisables combinant ICO et la RA ?

13. Un paradigme ICO est un type d'activité ou de réaction du cerveau qui est exploité pour l'interaction

Jusqu'au jour où ces lignes sont écrites, il est impossible de décoder les "pensées" et l'activité cérébrale, tel qu'imaginé par le grand public. Les ICO utilisent plutôt des propriétés neurophysiologiques documentées et des réactions du cerveau connues, pour déduire les états mentaux et les commandes souhaitées. Le choix d'une propriété plutôt que d'autres revient au concepteur, compte tenu de la tâche pour laquelle l'ICO est conçue. Ensuite, pour transformer l'activité du cerveau en une commande informatique, une ICO suit une série d'étapes résumées dans la Figure 8.9. Lorsque la propriété neurophysiologique souhaitée a été choisie, la première étape consiste à *mesurer* l'activité du cerveau. Pour cela, plusieurs techniques d'imagerie cérébrale ont été développées au fil des années. La plus utilisée cependant, et aussi la plus ancienne, est appelée Electroencéphalographie (EEG) [Berger, 1929]. L'EEG mesure l'activité électrique du cerveau au moyen d'électrodes placées dans différentes régions du cuir chevelu, et donne un signal représentant son évolution temporelle. Cependant, ce signal obtenu est toujours contaminé par du bruit. L'étape suivante, appelée *prétraitement*, consiste donc à séparer ce bruit du signal utile. L'étape suivante, appelée *extraction de caractéristiques*, consiste à extraire les informations utiles du signal prétraité. Cette étape a 2 objectifs : (1) réduire la dimensionnalité du signal et (2) décrire le signal avec des caractéristiques significatives. Ces caractéristiques sont ensuite *classifiées* dans des états mentaux qui sont ensuite traduits en commandes qui sont envoyées au système informatique. Tout ce processus, également connu sous le nom de "*Online BCI processing pipeline*" est souvent complété par une étape de rétroaction pour aider l'utilisateur à ajuster la modulation de son activité cérébrale.

D'autre part, la Réalité Augmentée est définie comme l'intégration d'objets et d'informations virtuels dans le monde réel en temps réel [Zhou et al., 2008]. Milgram et Kishino ont établi un continuum allant de la virtualité complète à la réalité complète. Entre eux, existent différentes combinaisons d'environnements réels et virtuels [Milgram and Kishino, 1994], selon le niveau de chacun dans la scène (voir Figure 8.10).

D'après Azuma, trois caractéristiques définissent un système RA : (1) la combinaison de contenu réel et virtuel, (2) l'interaction en temps réel, (3) l'enregistrement 3D du contenu virtuel dans l'environnement réel. Contrairement à la réalité virtuelle (RV), où l'utilisateur est immergé dans un monde complètement virtuel, la RA mélange les contenus virtuels et réels, idéalement, en les faisant sembler coexister dans le même espace.

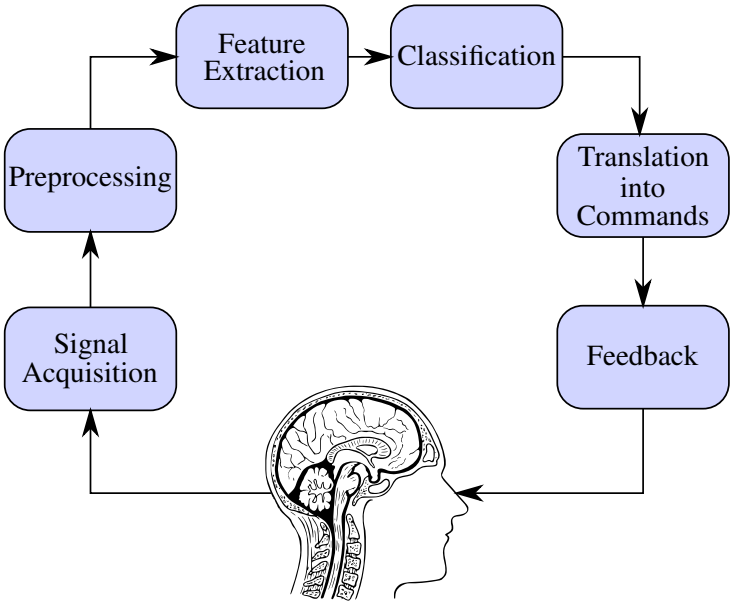


Figure 8.9: Illustration des étapes de conception d'une ICO. Ces étapes, telles qu'inspirées par [Mason and Birch, 2003], résument les différents composants d'une ICO.

Objectifs

Compte tenu de l'objectif global de ce travail, qui consiste à étudier la combinaison des BCI et de la RA, nous avons pu identifier cinq objectifs intermédiaires :

1. **Introduire une nouvelle perspective des ICO :**

   Plutôt que de considérer une ICO comme un simple dispositif de communication, il est important de comprendre ce qu'elle peut offrir comme source d'information sur l'état mental de l'utilisateur. À ce jour, il n'y a pas de consensus sur la définition du terme " interface cerveau-ordinateur ". Plusieurs définitions ont cependant été proposées, et elles ont tendance à refléter la diversité des domaines des chercheurs. La communauté médicale peut considérer les ICO comme des systèmes de réadaptation alors qu'un roboticien peut les percevoir comme des systèmes de contrôle de robots. Nous soutenons qu'une ICO peut être les deux. Le premier objectif de ce travail est de donner une nouvelle perspective sur la signification et la terminologie derrière le terme ICO. Avec cette nouvelle perspective, le but serait de mieux comprendre l'intégration de ces technologies dans la RA.

2. **Étudier la combinaison de l'ICO et de la RA en fonction des paradigmes de l'ICO :**

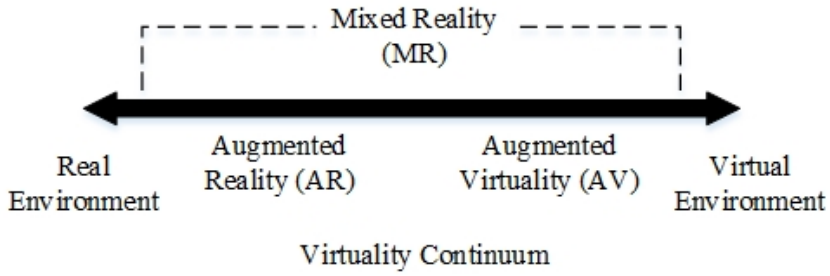


Figure 8.10: Le "virtuality continuum" de Milgram and Kishino [Milgram and Kishino, 1994] représentant les différents niveaux de mélange d'éléments réels et virtuels.

Le deuxième objectif de ce travail est d'étudier l'utilisation de différents paradigmes des ICO en RA. En particulier parce que la combinaison d'ICO et de la RA peut générer des interférences car l'EEG est très sensible au bruit [McFarland et al., 2005], il est important d'évaluer la possibilité d'utiliser tous les paradigmes en RA. De plus, quels types d'activité cérébrale est-il possible de détecter lorsqu'on utilise une ICO en conjonction avec la RA ? Et quels types d'événements ou de stimuli dans la RA peuvent susciter des réponses cérébrales aussi spécifiques ? Le deuxième objectif de ce travail est donc d'étudier les effets de la combinaison d'ICO et de la RA par rapport à divers paradigmes ICO.

3. **Améliorer l'ICO elle-même :**

   L'une des principales limites des ICO actuellement, est leur manque de précision [Lotte et al., 2018]. En particulier, une des difficultés des ICO réside dans la reconnaissance des états de repos. Il est difficile de concevoir des systèmes capables de détecter la volonté et l'absence de volonté d'interaction des utilisateurs. Ces systèmes, appelés asynchrones ou à rythmes propres, sont une exigence pour tout système interactif pratique comme dans la RA. Ainsi, le troisième objectif de ce travail est de proposer de nouvelles méthodes pour améliorer la précision de détection des ICO à rythme propre afin d'améliorer l'interaction dans la RA à l'aide d'une ICO.

4. **Conception de nouvelles interfaces pour l'utilisation d'ICO en RA :**

   La conception d'un système interactif comporte deux dimensions : la *performance*, c'est-à-dire la qualité du fonctionnement du système, et l'*intuitivité* de son interface, c'est-à-dire son caractère intuitif et sa facilité d'utilisation et de compréhension. En plus de la performance d'un système combinant RA et ICO, notre quatrième objectif est de caractériser et d'évaluer l'effet de l'interface utilisateur (IU) en RA sur la perception générale des utilisateurs et la performance du système.

5. **Développer de nouveaux usages de la BCI et de la RA :**

   Enfin, notre cinquième objectif est d'étudier les avantages de la combinaison de la RA et de l'ICB au moyen de cas d'utilisation concrets.

Approche et Contributions

Ce manuscrit décrit le travail et les contributions qui ont été réalisés afin d'atteindre les objectifs mis en évidence précédemment. Il est divisé en quatre parties : La **Partie I** donne un aperçu de la technologie des ICO ainsi qu'un état de l'art de la combinaison des ICO et de la RA. La **Partie II** décrit nos contributions pour améliorer la compréhension et la définition des ICO. La **Partie III** décrit nos études et nos contributions concernant l'évaluation des paradigmes des ICO qu'il est possible d'utiliser en RA, et la **Partie IV** détaille nos contributions concernant l'utilisation des ICO pour l'interaction active. [14] en RA. Un aperçu détaillé est donné ci-après :

Parti I : Etat de l'art

- **Chapitre 1: Introduction aux interfaces cerveau-ordinateur**
  Le but du premier chapitre est de donner au lecteur un aperçu de la technologie ICO. En partant des notions de base sur le cerveau, nous décrivons les paradigmes d'ICO les plus utilisés. Nous passons en revue le pipeline de conception des ICOs et présentons les méthodes et les algorithmes utilisés pour traduire l'activité cérébrale en entrées informatiques. Ce chapitre présente également une taxonomie des différents types de BCI.
- **Chapitre 2: Interfaces cerveau-ordinateur et réalité augmentée**
  Le but du deuxième chapitre est de faire le point sur l'état de l'art de l'utilisation des interfaces cerveau-ordinateur (ICO) en combinaison avec la réalité augmentée (RA). Il présente d'abord le domaine de la RA et ses principaux concepts. Ensuite, il décrit les différents systèmes combinant la RA et les BCI dans la littérature, classés selon leur domaine d'application : médecine, robotique, domotique et visualisation de l'activité cérébrale. Enfin, il résume et discute les résultats de l'enquête.

14. Les ICO pour l'interaction active désigne les scénarii où l'ICO est utilisée pour envoyer activement des commandes.

Partie II : Sur la définition des interfaces cerveau-ordinateur

Dans cette partie, nous visons à introduire une nouvelle perspective sur la définition d'une BCI.

- **Chapitre 3 : Proposition d'une nouvelle définition des ICO d'un point de vue de interaction homme-machine**
  Dans le troisième chapitre, nous présentons une nouvelle définition des interfaces cerveau-ordinateur. Du point de vue de l'interaction homme-machine, nous soutenons qu'une ICO doit être considérée comme une *Interface* reliant un utilisateur, par son activité *Cerebrale*, à tout système informatique *Ordinateur*. Nous discutons des définitions proposées précédemment, et nous établissons une analogie entre une BCI et une souris, une Interface Homme-Ordinateur répandue.

Partie III : Paradigmes ICO en RA

Cette partie rassemble nos contributions concernant l'essai des différents paradigmes ICO en RA. Notre objectif est d'évaluer la possibilité (1) d'utiliser le paradigme des potentiels évoqués visuels en état d'équilibre (SSVEP) pour l'interaction directe dans la RA, et (2) de susciter et de détecter les potentiels liés aux erreurs (ErrP) dans la RA.

- **Chapitre 4 : Utiliser SSVEP pour l'interaction ICO-RA**
  Dans le quatrième chapitre, nous étudions l'utilisation du paradigme SSVEP dans la RA. Nous étudions l'effet du contexte de la RA sur la possibilité de susciter et de détecter les réponses SSVEP. Ensuite, comme la RA est rarement statique, et qu'elle implique souvent des scénarii où l'utilisateur est libre de ses mouvements, nous évaluons l'effet du mouvement de la tête sur les réponses de la SSVEP en RA.
- **Chapitre 5 : Vers la détection des erreurs système dans la RA à l'aide des ErrPs**
  Dans le cinquième chapitre, nous étudions la possibilité d'obtenir et de détecter les potentiels liés aux erreurs (ErrP) dans la RA. En particulier, nous étudions l'effet des erreurs courantes dans les systèmes de RA (erreurs de suivi, erreurs de rétroaction et anomalies de fond) sur les signaux EEG des utilisateurs et la possibilité de détecter automatiquement les ErrP provoqués.

Partie IV : Utilisation de la BCI basée sur la SSVEP pour interagir dans la RA

Cette partie rassemble nos contributions concernant l'utilisation du paradigme SSVEP pour interagir en RA. Suite à l'essai de deux paradigmes ICO (SSVEP et ErrP) en RA (Partie III), nous avons choisi de concentrer nos efforts sur le paradigme SSVEP, car il s'agit de l'un des paradigmes ICO basés sur l'EEG les plus prometteurs [Chen et al., 2015]. Notre approche a consisté en 3 étapes : (1) améliorer la performance de l'ICO basée sur SSVEP, elle-même, (2) améliorer l'intuitivité de l'interface utilisateur SSVEP en RA, et (3) améliorer l'application en utilisant une ICO basée SSVEP en RA.

- **Chapitre 6: Amélioration de la classification asynchrone des SSVEP à l'aide de l'Analyse Canonique de Correlation (CCA) et des modèles de Markov cachés**
  Dans le sixième chapitre, nous visons à améliorer la précision de la classification des réponses SSVEP asynchrones afin d'améliorer l'interaction entre en utilisant SSVEP en RA. Pour ce faire, nous présentons *HCCA* une nouvelle méthode basée sur l'analyse de la dynamique des coefficients de corrélation canonique par des modèles de Markov cachés. La méthode est présentée en détails, et les résultats comparés à l'état de l'art, montrent que l'HCCA surpasse les méthodes existantes en termes de classification des états de non-contrôle, tout en affichant des taux de classification élevés des commandes.

- **Chapitre 7: Interfaces AR pour les ICO basées sur le SSVEP**
  Dans le septième chapitre, nous caractérisons l'espace de conception des stratégies d'affichage des interfaces utilisateur SSVEP dans la RA. Nous évaluons l'effet de ces stratégies d'affichage en termes de perception de l'utilisateur et de précision de la classification SSVEP, dans le but d'extraire des lignes directrices sur la façon de faire un compromis entre la performance et l'intuitivité du système.
- **Chapitre 8: Combinaison d'ICO et de RA pour la domotique**
  Dans le huitième chapitre, nous proposons des lignes directrices générales pour le développement de systèmes domotiques basés sur la RA et les ICO. Nous proposons des spécifications globales que ces systèmes devraient respecter et nous concevons une architecture générique pour aider au développement de systèmes maintenables. Nous décrivons également une validation de principe industrielle de l'intégration d'un système de RA et d'une ICO à une plate-forme domotique comerciallement disponible.

ETAT DE L'ART

Les interfaces cerveau-ordinateur et la réalité augmentée sont deux domaines qui peuvent être combinés à des fins d'interaction et/ou de visualisation. D'une part, les systèmes basés sur la RA reposent généralement sur des écrans montés sur la tête (HMD), qui peuvent être utilisés dans des scénarios nécessitant une interaction mains libres [Blum et al., 2012]. Les paradigmes ICO peuvent fournir de tels moyens d'entrée, soit pour interagir avec des objets virtuels [Faller et al., 2017] ou réels [Takano et al., 2011]. D'autre part, les ICO peuvent tirer parti de la RA afin de permettre l'interaction avec le monde réel. La RA peut également fournir des moyens intéressants d'afficher des retours utilisateurs en les intégrant dans l'environnement du monde réel.

La Réalité Augmentée peut généralement être obtenue de 3 manières différentes [Azuma, 1997, Azuma et al., 2001] (voir Figure 8.11) :

1. Vidéo See-Through (VST) AR : dans laquelle des images réelles sont filmées par la caméra d'un appareil (tablette, téléphone, hmd) avant d'être réaffichées à travers un écran, enrichies d'informations virtuelles ;
2. RA optique transparente (OST) : dans laquel le contenu virtuel est affiché directement devant les yeux de l'utilisateur sur un écran semi-transparent (p. ex., Microsoft Hololens)
3. RA projetée (aussi appelée réalité augmentée spatialement) : dans laquelle un contenu virtuel est projeté sur un objet de l'environnement réel

La combinaison de la RA et des ICO pourrait être appliquée à la plupart des sujets où les ICO peuvent être utilisées, par exemple l'assistance aux personnes handicapées, les loisirs, les sports. Il peut y avoir différentes raisons de combiner RA et ICO. Tout d'abord, du point de vue de la BCI, la RA offre de nouvelles façons d'intégrer le retour d'information dans l'environnement réel, ce qui apporte de nouvelles possibilités d'interaction et améliore l'expérience de l'utilisateur. Deuxièmement, du point de vue de la RA, les ICB offrent de nouveaux paradigmes mains libres pour l'interaction avec des objets physiques et virtuels ainsi que de nouvelles informations physiologiques sur l'état mental de l'utilisateur, ce qui permet de créer des scénarios plus adaptatifs.

15. Sources d'images: (Left) https://www.wired.com/2015/04/microsoft-build-hololens/ (Middle) https://www.youtube.com/watch?v=J8qDRV1RxTw

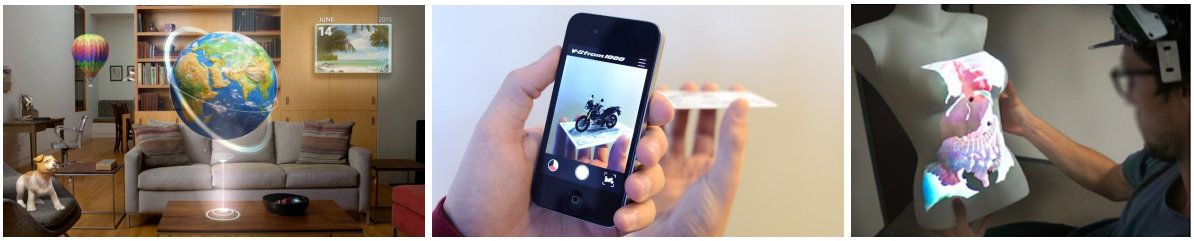


Figure 8.11: Différents types de RA [15]. (Gauche), éléments virtuels directement superposés dans le champ de vision de l'utilisateur (Optical See-Through AR). (Au milieu), un élément virtuel ajouté à un flux vidéo enregistré sur un téléphone (Vidéo See-Through). (A droite), éléments virtuels superposés sur un objet physique (Réalité Spatiale Augmentée) [Cortes et al., 2018]..

L'analyse de l'état de l'art montre d'abord que l'augmentation se fait la plupart du temps soit par des écrans d'ordinateur soit par des casques, et que seul un petit nombre d'entre eux utilisent la RA optique transparente. La raison est peut-être que la première solution est pratique pour le prototypage et la seconde très intuitive, permettant plus de mobilité pour les utilisateurs. Cependant, si la RA sur écran empêche clairement les utilisateurs de se déplacer, l'état de développement des BCI jusqu'à présent les empêche également de se déplacer avec les HMD en raison du risque d'artéfacts musculaires.

Comme la combinaison de la RA et des ICO est relativement nouvelle, la question de la mobilité ne semble pas avoir été abordée dans la plupart des articles utilisant des HMD. Cependant, le développement et l'amélioration de la technologie des ICB, notamment la mise au point de méthodes de filtrage pour éliminer efficacement les artéfacts musculaires, seraient une condition préalable à l'utilisation des ICB comme outil d'interaction avec la RA à son plein potentiel.

En ce qui concerne les paradigmes ICO, bien qu'un certain nombre d'entre eux aient été pris en compte, les paradigmes SSVEP et P300 sont les plus utilisés. Cette popularité pourrait être due à l'aspect graphique de l'augmentation, puisque la RA est basée sur l'affichage d'éléments graphiques virtuels sur le champ de vision de l'utilisateur, donc les paradigmes basés sur la vision sont bien adaptés aux tâches de sélection.

Cependant, il est important d'explorer plus en profondeur l'effet de la RA sur les performances des ICO, non seulement du point de vue des performances du système mais aussi en termes d'interfaces utilisateurs.

Il ressort également que la plupart des travaux s'appuient sur des paradigmes actifs de la BCI (y compris réactifs). Ils étaient surtout utilisés pour la manipulation et le contrôle volontaire d'objets physiques ou virtuels. Les BCI passives pourraient également être étudiées plus en

profondeur dans le contexte de la RA.

En tenant compte de ces informations, nous étudions dans ce travail la possibilité d'utiliser différents paradigmes BCI dans la RA, à savoir SSVEP et ErrPs. L'objectif étant d'évaluer la possibilité de les détecter dans les signaux EEG des utilisateurs en RA ainsi que d'évaluer l'effet des mouvements sur leur précision de détection. Nous considérons également le cas d'application prometteur de la combinaison de la BCI et de la RA : la domotique.

PROPOSITION D'UNE NOUVELLE DÉFINITION DES IC DU POINT DE VUE DE L'INTERACTION HOMME-MACHINE

En 1973, Jacques Vidal a introduit le terme Interface Cerveau-Ordinateur [Vidal, 1973]. Le projet "ICO" tel qu'il était imaginé à l'époque, était destiné à exploiter l'activité électrique provenant du cerveau pour contrôler un programme informatique. Même si d'autres termes peuvent être trouvés dans la littérature pour désigner les ICO : Interfaces Cerveau-Machines(ICM), Interfaces Cérébrales Directes (ICD), Interfaces Neuronales Directes (IND) ou Interfaces Cérébrales (IC). Il est admis qu'elles désignent toutes la même chose : une interface entre un cerveau et un ordinateur (au sens large du terme). Cela se remarque même par le fait que la plupart d'entre elles partagent le "I" de "Interface", terme couramment utilisé dans l'Interaction Homme-Ordinateur, le domaine qui vise à concevoir, évaluer et mettre en œuvre des systèmes interactifs [Hewett et al., 1992].

La définition la plus largement reconnue d'une ICO a été proposée par Wolpaw et al. en 2002 : "*Une interface cerveau-ordinateur est un système de communication dans lequel les messages ou les commandes qu'un individu envoie au monde extérieur ne passent pas par les voies de sortie normales du cerveau des nerfs et des muscles périphériques*" [Wolpaw et al., 2002]. Le principal objectif d'une ICO serait alors de permettre aux utilisateurs qui souffrent peut-être du syndrome d'" enfermement " ou de " paralysie " de communiquer et d'exprimer leurs souhaits aux soignants, ou d'utiliser des programmes de traitement de texte et des neuroprothèses. L'idée fondamentale d'une ICO serait donc de fournir au cerveau "un nouveau canal de communication et de contrôle non musculaire pour transmettre des messages et des commandes au monde extérieur". Cela fait apparaître un élément essentiel d'une ICO qui est la nature " non musculaire ". Une ICO tire son apport uniquement de l'activité cérébrale.

Dans le dictionnaire Oxford [16], le mot " interface " définit " *Un point où deux systèmes, sujets, organisations, etc. se rencontrent et interagissent* ". Dans le contexte de l'informatique, il

16. Interface In : Oxford University Press. [en ligne] Disponible sur : https://en.oxforddictionaries.com/definition/interface [consulté le 20 février 2019].

définit " *Un dispositif ou un programme permettant à un utilisateur de communiquer avec un ordinateur* ". De ces deux définitions, il ressort clairement que la notion d'interface n'a de sens que si l'on considère les éléments qu'elle permet de relier. Dans l'IHM, ces deux éléments sont l'"utilisateur" et le "système interactif".

Dans ce contexte, une interface est l'ensemble des moyens matériels et logiciels par lesquels l'utilisateur communique avec le système interactif. L'interface comprend le dispositif d'entrée (le matériel), les algorithmes et les méthodes pour traiter ses sorties et les mécanismes de présentation pour afficher le retour d'information. L'objectif final du système, cependant, n'appartient pas aux frontières de l'interface.

Selon ce principe, un ICO peut être considérée comme n'importe quelle interface homme-machine, puisqu'il comprend le même ensemble de composants. Par conséquent, dans la perspective d'une ICO, nous proposons de définir une interface cerveau-ordinateur comme suit :

Définition

Une ICO est tout système artificiel qui transforme directement l'activité cérébrale en entrée d'un processus informatique.

Cette définition présente l'avantage d'être inclusive car elle ne comprend pas la finalité du système interactif. Une réflexion similaire a été menée par Jeunet et al. [Jeunet et al., 2018] en comparant le Neurofeedback et l'Imagerie Motrice. Ils ont souligné que les utilisateurs de Neurofeedback (NF) et de MI-BCI doivent apprendre à réguler leur activité neurophysiologique, parfois avec des caractéristiques similaires, par un feedback donné avec un objectif final différent. Alors que le MI-BCI consiste à produire un schéma EEG spécifique pour envoyer une commande, le but du NF est d'apprendre à générer ce schéma spécifique.

Si l'on veut désigner l'ensemble de l'interaction entre un utilisateur et un système interactif utilisant une BCI pour remplir un rôle particulier, on devrait utiliser le terme d'interaction cerveau-ordinateur.

Avec notre nouvelle définition des interfaces cerveau-ordinateur (BCI), tout système artificiel qui implique l'exploitation d'un signal cérébral peut être considéré comme utilisant une BCI, au lieu d'être la BCI elle-même.

SSVEP : ÉTUDE ET ÉVALUATION DES RÉPONSES SSVEP EN RA

Introduction

La SSVEP est un schéma cérébral spécifique qui se produit lorsque le système visuel humain est stimulé par un clignotement périodique. Le cerveau répond par une activité à la même fréquence dans la zone corticale visuelle [Herrmann, 2001]. Lorsque l'utilisateur est confronté à plusieurs cibles clignotant à des fréquences différentes, il devient possible de déterminer sur quelle cible il se concentre en analysant les signaux de son électroencéphalographie (EEG).

L'objectif de ce chapitre est donc d'étudier la faisabilité de combiner le SSVEP et la RA. Nous avons mené deux études auprès des utilisateurs, en évaluant les baisses potentielles de la performance de l'ICB dues à :

1. Le port d'un casque EEG et d'un casque AR.
2. Les mouvements de la tête de l'utilisateur lors de l'observation de la scène AR.

Étude utilisateurs : Influence du dispositif de RA sur les performances du BCI

L'objectif de la première étude était de répondre aux questions suivantes :

1. La qualité du signal EEG est-elle altérée par le port d'un OST-HMD ?
2. Est-ce que l'activité électrique produite par l'écran corrompt les signaux ?
3. Le paradigme de la SSVEP est-il efficace dans un tel contexte de RA stéréoscopique ?

Ces questions sont évaluées en comparant les performances de la BCI tout en interagissant avec une interface SSVEP dans différentes configurations matérielles. Treize participants naïfs (25,9 ans, 3,1 ans, 2 femmes) ont pris part à l'expérience.

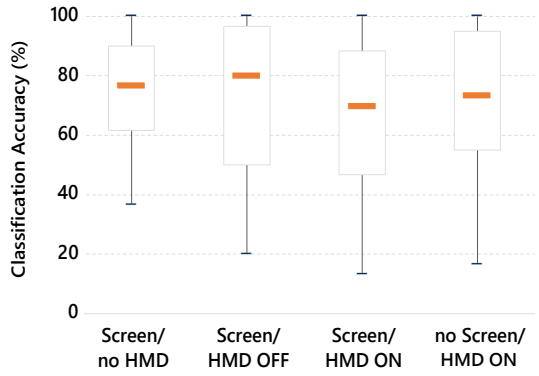


Figure 8.12: Résultats de l'étude 1 : Boxplots représentant la précision de la classification en % (N=13) en fonction de la condition visuelle.

Protocole expérimental

Chaque session d'expérimentation a duré environ 7 minutes (30 essais) et consistait en l'affichage de 3 cibles blanches soit à l'écran soit sur HoloLens. Chacune de ces cibles clignotait à une fréquence différente (10, 12 et 15 Hz).

Résultats

Les résultats sont résumés dans la Figure 8.12.

La précision de la classification dans toutes les conditions était comprise entre 75 % et 80 %, ce qui est conforme à la littérature [Friman et al., 2007, Chen et al., 2014].

Étude utilisateurs : influence des mouvements des utilisateurs sur la performance de la BCI en matière de RA

L'objectif de cette deuxième étude était d'évaluer l'impact des mouvements de la tête (source de bruit interne) sur la performance de sélection de l'ICO afin de répondre aux questions :

1. Est-il possible d'effectuer une sélection SSVEP sur des cibles mobiles dans la RA ?
2. Le mouvement de la tête influence-t-il la performance de l'ICB ?

Afin d'évaluer la possibilité d'utiliser des ICO basées sur la SSVEP sur des cibles mobiles, nous avons comparé la performance des utilisateurs sur la sélection des cibles pour différentes

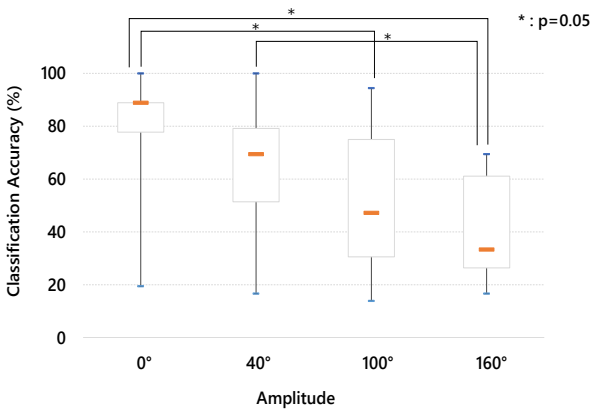


Figure 8.13: Résultats de l'étude 2 : Boxplots représentant la précision de la classification en % (N=15) en fonction de l'intensité du mouvement.

intensités de mouvement. Quinze participants (âgés de 22,8 ans, 1 femme) ont pris part à l'expérience.

Chaque Session a duré environ 7 minutes et comprenait 36 essais.

Résultats

Les résultats de l'expérience montrent une nette dégradation de la performance de l'ICB ainsi qu'une augmentation de l'intensité des mouvements (voir Figure 8.13). Ainsi, la performance moyenne chute de 78%, lorsque les participants ne bougent pas, à 41% (ce qui se situe autour du niveau de chance [Müller-Putz et al., 2008a]) pour l'amplitude maximale des mouvements.

Cependant, après avoir effectué un test post-hoc par le biais de comparaisons multiples entre les conditions il est apparu que la différence entre les conditions "Pas de mouvement" et "Mouvement de faible intensité" n'était pas significative (voir figure 8.13).

Conclusion

Dans ce chapitre, nous avons étudié la combinaison des interfaces cerveau-ordinateur basée sur le paradigme SSVEP et la réalité augmentée. Nous avons constaté que l'utilisation d'un casque d'écoute AR n'a pas altéré de façon significative la performance de l'ICB et n'a pas perturbé la précision de la classification EEG. Il semble donc possible d'exploiter les données SSVEP et EEG avec un casque AR sur le capuchon EEG. Nous avons également constaté qu'il est possible d'utiliser une ICB basée sur la SSVEP en présence de petits mouvements de la tête.

ErrPs : Vers la détection des erreurs de système en RA à l'aide des ErrPs

Ce chapitre vise à étudier la présence et la détection des Error Related Potentials (ErrPs) en Réalité Augmentée. En particulier, nous étudions la présence et la détection des ErrPs lorsque les utilisateurs sont confrontés à des erreurs système, à savoir : les erreurs de suivi, les erreurs de résultat et les anomalies de fond. L'objectif est d'exploiter la robustesse et la contrôlabilité des environnements de Réalité Virtuelle (RV) pour simuler les erreurs de RA, afin de mener les expériences dans des conditions mieux contrôlées.

Une ErrP est déclenchée lorsque l'utilisateur commet une erreur ou est témoin d'une erreur. Elle peut être observée dans les signaux EEG de l'utilisateur sous la forme d'un schéma EEG spécifique apparaissant peu après le début d'un événement erroné.

Matériel et Méthodes

Nous avons identifié 3 types de situations susceptibles de provoquer des potentiels d'erreurs : (Te) La perte de suivi d'un objet manipulé ; (Fe) Un résultat inattendue ou erronée ; (Be) Une anomalie de fond irréaliste. D'après la littérature et les travaux précédents, nos hypothèses étaient que (Te) provoquerait une Négativité Liée aux Événements (NRE), associée aux erreurs d'exécution, que (Fe) déclencherait une Négativité Liée à la Rétroaction tandis que (Be) déclencherait une NRE pour les utilisateurs qui remarquent l'anomalie.

Appareil et participants

Quinze participants en bonne santé et non daltoniens [17] (7 femmes, 8 hommes) ont pris part à l'expérience (moyenne d'âge = 24,8 ans, std = 2,9).

17. Cette recherche a été menée conformément aux directives pertinentes pour la recherche éthique selon la Déclaration d'Helsinki. Tous les participants ont été informés de la nature de l'expérience et ont signé un formulaire de consentement éclairé au début de l'expérience.

Figure 8.14: Illustration de 4 conditions expérimentales. En haut à gauche : *Correct Feedback (Fc)* condition. Le participant reçoit un résultat correct après l'achèvement de la tâche. En bas à gauche : *Erreur de résultat (Fe)* condition. Le participant reçoit un résultat incorrect après l'achèvement de la tâche. En haut à droite : *Feedback error (Fe)* condition : *Erreur de suivi (Te)* condition. L'objet se fige et se détache de la main du participant. En bas à droite : *Anomalie de fond (Be)*. Le cadre devant le participant se retourne et se met dans une position irréaliste, pénétrant la table.

L'expérience consistait en 4 conditions correspondant aux types d'essais :

- **Correct:** Ces essais correspondent à la réalisation normale de la tâche. Dans ces essais, le participant a saisi l'objet et l'a laissé tomber dans la forme correcte. Il a reçu le feedback correct (Fc) (Figure 8.14 : En haut à gauche).
- **Erreur de suivi (Te):** Ces essais correspondent à la situation dans laquelle le système a perdu le suivi de l'objet avant que le participant ne le place dans la forme cible. L'objet s'est figé et s'est détaché de la main du participant. (Figure 8.14 : En haut à droite).
- **Erreur de rétroaction (Fe):** Ces essais correspondent au fait que le participant reçoit un résultat inattendu et erroné après avoir terminé la tâche. Même si l'objet était placé dans la forme correcte, la forme deviendrait *red*, ce qui correspond à un feedback erroné (Figure 8.14 : En bas à gauche).
- **Anomalie de fond (Be):** Ces essais correspondent à la situation où une anomalie irréaliste, non liée à la tâche, apparaît en arrière-plan. Un cadre placé devant le participant sur la table se retournait et restait dans une position irréaliste (traversant la table) jusqu'à la fin de l'essai (voir Figure 8.14 : En bas à droite).

Résultats

Analyse électrophysiologique

Analyse de l'erreur de suivi vs le suivi correct (Te et Tc)

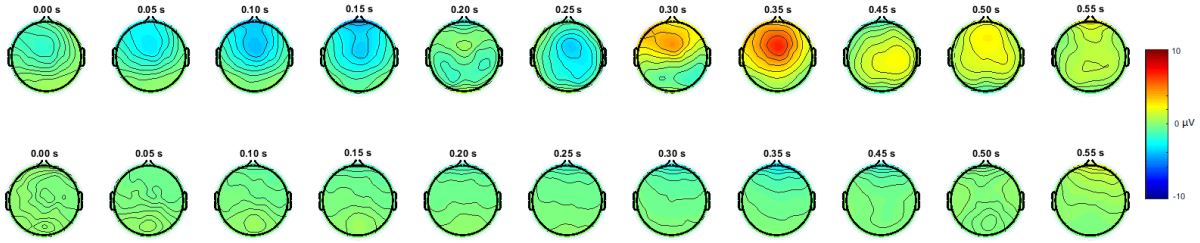


Figure 8.15a: Toplots de la moyenne générale des conditions d'erreur (Te) et des conditions correctes (Tc) (lignes supérieure et inférieure respectivement), affichés de $t = 0$ s à $t = 0,55$ s par intervalles de 50 ms.

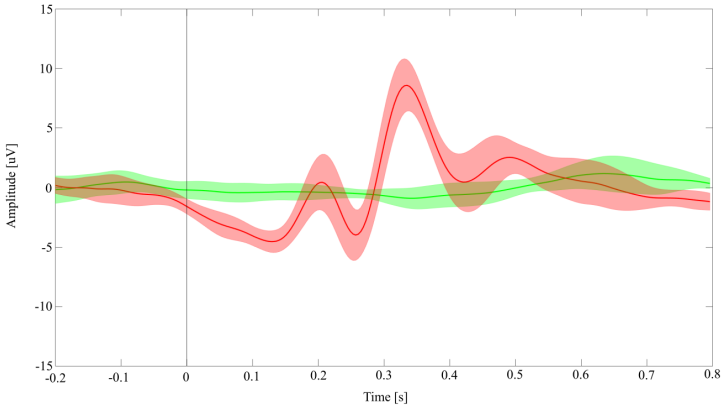


Figure 8.15b: Moyenne générale des signaux corrects (Tc) et d'erreur (Te) sur le canal FCz (lignes verte et rouge, respectivement). Les zones ombrées représentent l'intervalle de confiance à 95 % pour les courbes moyennes. La ligne verticale noire représente les événements de début.

Les résultats de l'analyse électrophysiologique des conditions de suivi (Te et Tc) sont résumés dans la figure 8.15.

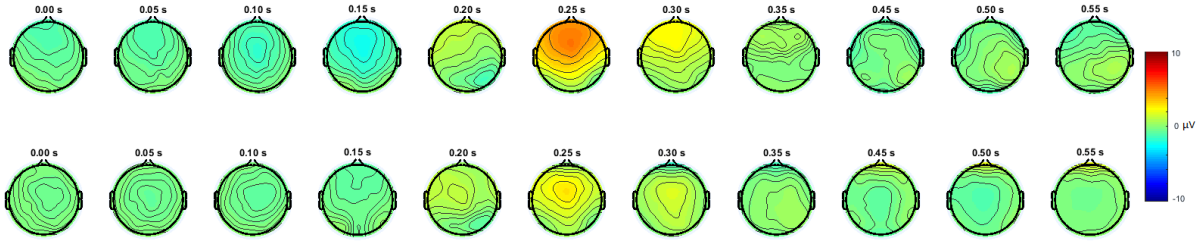


Figure 8.16a: Toplots de la moyenne générale des conditions d'erreur (Fe) et des conditions correctes (Fc) (lignes supérieure et inférieure, respectivement), affichés de $t = 0\,\text{s}$ à $t = 0,55\,\text{s}$ par intervalles de 50 ms.

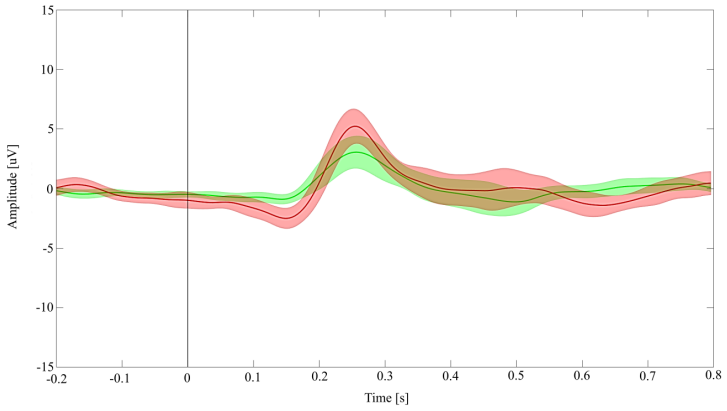


Figure 8.16b: Moyenne générale des signaux corrects(Fc) et d'erreur (Fe) sur le canal FCz (lignes verte et rouge, respectivement). Les zones ombrées représentent l'intervalle de confiance à 95 % pour les courbes moyennes. La ligne verticale noire représente les événements de début.

Analyse de l'erreur de résultat par rapport à un résultat correct (Fe et Fc)

Les résultats de l'analyse électrophysiologique des conditions de rétroaction (Fe et Fc) sont présentés sur la figure 5.5.

Classification de l'essai unique

Afin d'illustrer la faisabilité de la détection des potentiels d'erreur dans un essai unique dans une expérience de RV, nous avons décidé de tester la classification de l'essai unique sur la condition de poursuite (Te vs Tc) et sur la condition de rétroaction (Fe vs Fc). Les signaux EEG concernant la condition de poursuite étaient très distincts et la condition Te présentait un schéma ErrP clair.

Condition de suivi

Le tableau 5.1 affiche les résultats de classification des essais Tc vs Te en termes de précision, TNR et TPR. La précision moyenne obtenue était de $84,9\%$. Le TNR moyen obtenu était de $88,4\%$. Le TPR moyen obtenu était de $81,5\%$. Les résultats obtenus sont significativement supérieurs au niveau de hasard ($alpha = 0,05$).

Feedback condition

Le tableau 5.2 affiche les résultats de classification des essais de Fc vs Fe en termes de précision, TNR et TPR. La précision moyenne obtenue était de $59,3\%$. Le TNR moyen obtenu était de $59,0\%$. Le TPR moyen obtenu était de $59,6\%$. Les résultats moyens ne sont pas significativement supérieurs au niveau de hasard ($\alpha = 0.05$).

Conclusion

Dans ce chapitre, nous avons étudié la possibilité de détecter les potentiels d'erreur dans la réalité virtuelle lorsque le système génère une erreur inattendue. En particulier, nous avons étudié trois types d'erreurs : les erreurs de suivi, les erreurs de rétroaction et les anomalies de fond. Les différentes erreurs représentaient un aspect différent sur le processus d'interaction (tâche, résultat, contexte) et avaient un impact différent sur la réalisation de la tâche.

Les résultats ont indiqué que seules les erreurs qui avaient un impact sur la tâche (pertes de suivi) étaient capables de générer des ErrPs puisqu'un schéma clair d'ErrP ne pouvait être observé que dans la Figure 5.5b De plus, nous avons montré que la détection des ErrPs liées aux erreurs de suivi dans la RV était possible avec une grande précision. Ceci est montré dans le tableau 5.1 où les erreurs de suivi ont été détectées de manière significative au-dessus du niveau de hasard pour tous les sujets.

NOUVELLE MÉTHODE POUR L'AMÉLIORATION DES ICO BASÉES SUR LE SSVEP : HCCA

Introduction

Dans ce chapitre, nous proposons d'utiliser les modèles de Markov cachés (HMM) [Rabiner, 1989] avec les coefficients de CCA comme vecteurs de caractéristiques afin de discriminer entre les états de non contrôle et de contrôle ainsi que d'effectuer une reconnaissance SSVEP. Nous avons appelé cette méthode HCCA.

La méthode proposée

HCCA repose sur l'hypothèse que lorsqu'un utilisateur passe d'un état d'absence de contrôle ou de repos à une commande clignotante, le coefficient CCA associé à la fréquence de stimulation a tendance à augmenter avec le temps. Ainsi, l'exploitation de ce comportement dynamique à l'aide des HMM permettrait d'améliorer la précision de la reconnaissance des SSVEP auto-régulées.

Pour résumer l'approche, les essais EEG correspondant à des " commandes SSVEP " ou " sans contrôle " sont découpés en segments qui se chevauchent. Chacun de ces segments est décrit par ses coefficients CCA, c'est-à-dire la corrélation canonique du segment avec les signaux de référence correspondant à chacune des fréquences de stimulation. Ce procédé permet d'obtenir un vecteur caractéristique pour chaque époque de l'essai. La concaténation des vecteurs caractéristiques correspondant aux Coefficients CC de ces époques successives, décrit la dynamique (évolution dans le temps) des Coefficients CCA.

Les séquences de vecteurs caractéristiques de chaque classe (correspondant à l'état "non-contrôle" et aux fréquences de stimulation) sont ensuite utilisées pour entraîner les HMM. Un HMM est formé pour chaque classe.

En mode fonctionnement, pour chaque nouvel essai EEG (découpé en segments et décrit à

l'aide des coefficients de CCA), la probabilité d'avoir été généré par chacun des HMM est estimée. La classification est ensuite réalisée selon un critère de maximum. La classe correspondant au HMM ayant la plus grande probabilité est désignée comme la classe de l'essai EEG. Toute cette approche est résumée dans la Figure 8.17.

Évaluation de la méthode

HCCA a été comparée et testée sur une population de 22 sujets ayant pris part à l'expérience. Chaque essai enregistré durait 5 secondes, et a été découpé en 2 segments de 2 secondes, se chevauchant sur 1s (4 segments par essai).

Nous avons effectué une validation croisée de 5-*groupes* des données enregistrées pendant l'expérience, en nous entraînant sur 4 cinquièmes des essais et en testant sur le cinquième restant.

Afin d'évaluer la performance des méthodes considérées, trois mesures de performance ont été utilisées (*Vrai Taux Positif*, *Vrai Taux Négatif* et *Précision de la Classification*) :

Résultats

Les matrices de confusion pour les groupes A, B et C utilisant l'ACCH sont données dans les tableaux 6.2, 6.3,et 6.4 respectivement.

Les matrices de confusion pour les groupes A, B et C utilisant le CCCA sont données dans les tableaux 6.5, 6.6 et 6.7 respectivement.

Comparaison des performances

Les résultats par sujet et les moyennes pour les groupes A, B, C sont résumés dans les tableaux 6.8, 6.9 et 6.10.

L'analyse statistiques de ces résultats montre que les participants du groupe C ont eu un TPR plus élevé avec la méthode HCCA (M = 69,17 ; $T = 11,78$) comparativement à la méthode CCCA (M = 53,17 ; $T = 2,48$). Cet effet n'a pas été observé pour les groupes A et B.

Enfin, pour ce qui est de la précision, nous n'avons pas trouvé d'effet significatif sur le groupe. Dans l'ensemble, les deux méthodes avaient un taux de précision comparable $M = 90,19; SD = 13,02$.

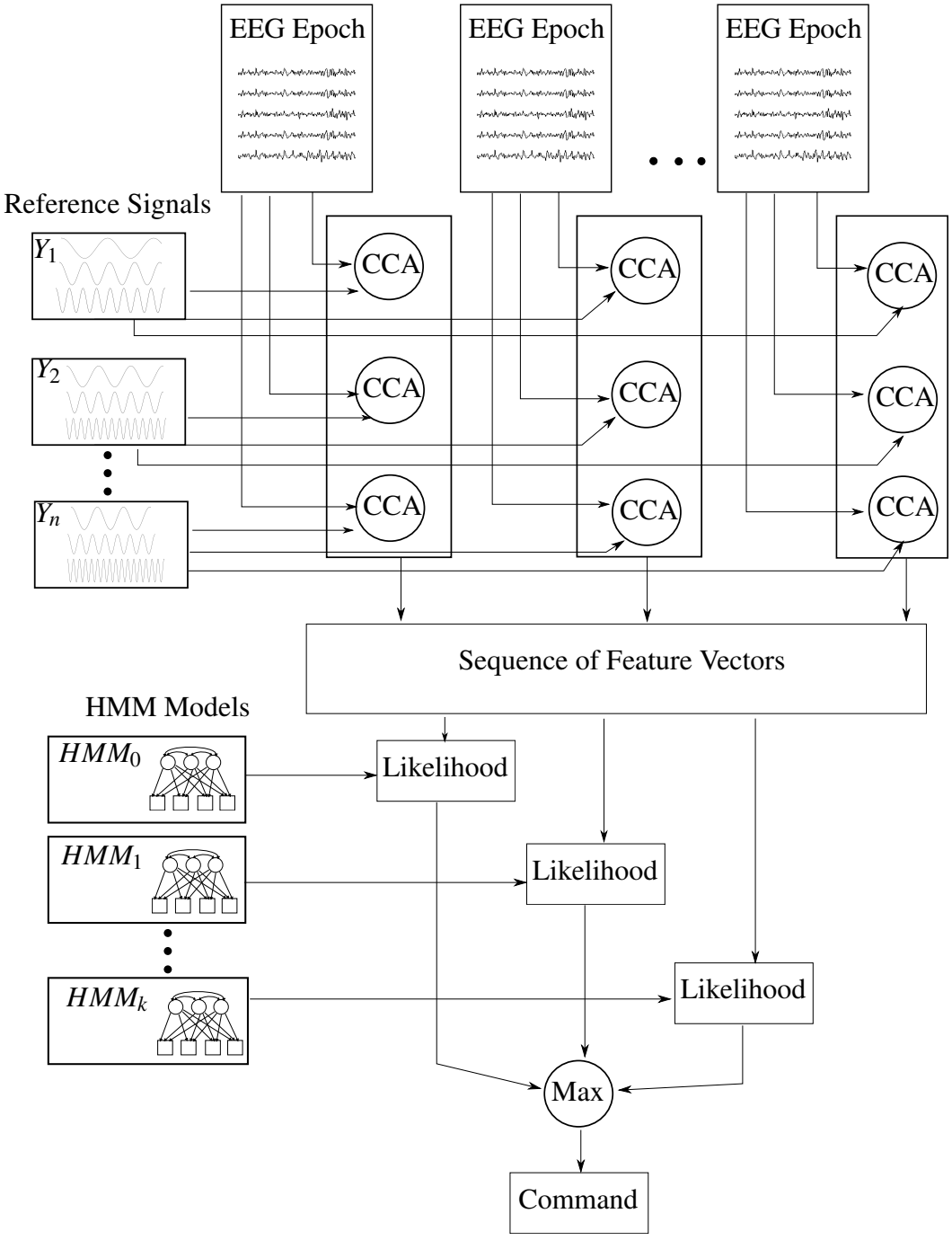


Figure 8.17: Représentation globale de la méthode proposée HCCA. Après la phase d'apprentissage pour estimer les paramètres HMM, l'analyse de corrélation canonique est calculée entre les époques successives constituant un essai et chaque signal de référence pour déterminer les Coefficients CC. Ces derniers sont agrégés dans une séquence de vecteurs caractéristiques représentant leur dynamique. Les probabilités que chaque modèle HMM entraîné ait généré le vecteur séquence sont calculées et un critère de maximum est appliqué pour déterminer la commande correspondante.

Conclusion

Nous avons présenté ici, une nouvelle méthode de reconnaissance des SSVEP asynchrones appelée HCCA. La HCCA utilise la dynamique CCA au cours du temps, lorsque les utilisateurs portent leur attention sur une stimulation de la SSVEP, à travers des modèles de Markov cachés qui permettent de modéliser et de classifier ces dynamiques. L'analyse de corrélation canonique a été utilisée pour l'extraction des caractéristiques, et les premières composantes de corrélation canonique avec les 3 fréquences de scintillement ont été agrégées comme vecteurs de caractéristiques. La séquence des vecteurs caractéristiques, représentant la dynamique des coefficients CC dans le temps, a été classée à l'aide de modèles de Markov cachés. Un modèle de Markov caché a été formé pour caractériser chaque classe de sélection (chaque fréquence de stimulation), ainsi que l'état de non contrôle. Chaque nouvelle séquence de vecteurs de caractéristiques a ensuite été classée selon le HMM qui était le plus susceptible de l'avoir générée.

L'étude utilisateur menée afin d'évaluer l'ACCH et de la comparer à l'état de l'art, a porté sur quatre conditions de la SSVEP : 3 cibles vacillantes et un point de fixation central. Ce dernier devait représenter l'état de non-sélection, où un utilisateur aurait des cibles clignotantes dans son champ de vision, sans vouloir envoyer une commande.

Les résultats de l'étude ont montré l'efficacité de l'HCCA pour la détection asynchrone des commandes SSVEP. Cette approche a été évaluée et comparée à l'analyse en grappes des coefficients CC, qui s'est avérée être la meilleure méthode dans l'état de l'art, et des tests statistiques ont montré que l'ACCH caractérise mieux les états de sélection et de non-sélection, en maintenant un meilleur taux de vrais négatifs.

Pris ensemble, ces résultats suggèrent que l'ACCH est une technique prometteuse pour l'interaction à rythme libre utilisant la SSVEP dans la RA. À condition qu'elle soit adaptée pour fonctionner en ligne.

CONCEPTION D'INTERFACES DE RA POUR LES ICO BASÉES SUR LA SSVEP

Introduction

L'objectif de ce chapitre est de proposer une méthodologie générique pour concevoir des interfaces utilisateurs pour les ICO basées sur le SSVEP. Nous avons considéré, comme cas d'utilisation de départ, le contexte de la commande d'un robot, en utilisant 3 commandes. Nous présentons un espace de conception étendu de différents agencements possibles, pour afficher 3 cibles pour la commande de robot basée sur SSVEP.

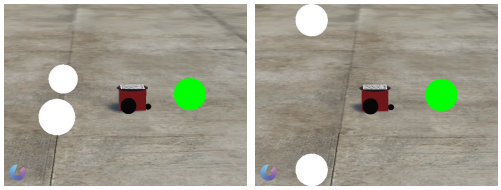
Pour illustrer le développement d'un prototype d'application basée sur la ICO en RA. Nous avons pu intégrer tous nos résultats précédents dans un dispositif unique dédié au contrôle d'un véritable robot mobile en RA au moyen d'une ICO basée sur SSVEP.

Espace de conception des stratégies d'affichage des cibles du SSVEP

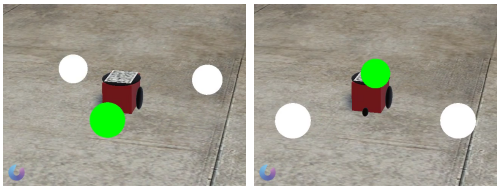
Afin de proposer l'association la plus cohérente de chaque commande avec la cible AR, nous proposons un espace de conception en 5 dimensions qui décrit toutes les dispositions possibles des 3 cibles de commande. Ces 5 dimensions, à savoir *Orientation, Cadre de référence, Ancrage, Taille et Explicitabilité*, sont présentées ci-dessous :

- **Orientation**: Cette dimension correspond à l'orientation du plan contenant les 3 cibles. Ce plan peut être "Transversal" c'est-à-dire parallèle à l'horizon ou "Frontal" c'est-à-dire orienté vers l'utilisateur (voir Figure **??**).
- **Référentiel**: Cette dimension correspond au référentiel des 3 cibles. Il peut être fixé au robot ou à l'utilisateur. Il peut également être associée aux concepts de référentiels "égocentrique" et "excentrique". (voir Figure **??**).
- **Ancrage**: Cette dimension correspond à la position du référientiel des cibles en 3D. Les positions peuvent être soit "ancrées" au robot, soit à l'utilisateur (voir Figure **??**).

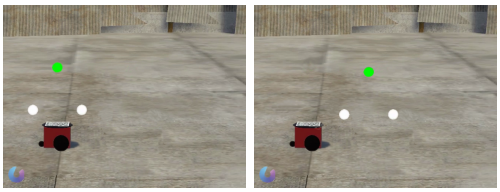
– **Taille**: Cette dimension est liée à la taille des cibles, qui peut également être un paramètre influent pour l'ergonomie du système. La taille des cibles peut être soit "Absolue" ou "Adaptative". Comme les cibles sont des objets 3D, nous avons utilisé cette terminologie pour définir leur taille réelle. Lorsque la taille est définie en absolu, le rayon de la cible est fixé dans l'espace 3D. Dans la condition adaptative, la taille réelle des positions dans l'espace 3D n'est pas fixée. Elle s'adapterait pour être toujours affichée à la même taille par rapport à l'utilisateur (voir Figure **??**).
– **Explicité**: Plus qu'une configuration d'affichage, cette dimension correspond à l'explicitation de l'association entre les cibles et les commandes. Une façon d'améliorer la sémantique de l'application est de lier explicitement une cible avec la direction vers laquelle elle mène. Lorsque le caractère explicite est défini sur "Explicite", la cible est reliée à la roue correspondante du robot (voir Figure **??**).



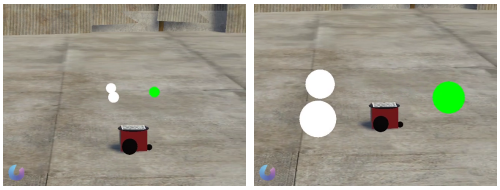
(a) Orientation: Transversal (left), Frontal (right)



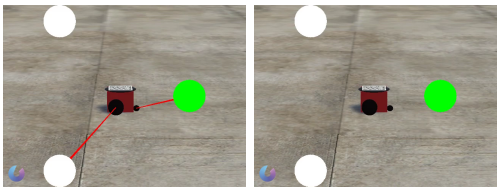
(b) Frame of Reference: Exo-Centered (left), Ego-Centered (right)



(c) Anchorage: Robot (left), User (Right)



(d) Size: Absolute (left), Adaptive (Right)



(e) Explicitness: Explicit (Left), Implicit (Right)

Figure 8.18: Illustration des 5 dimensions de notre espace de conception des cibles SSVEP : stratégies d'aménagement et d'affichage. Chaque couple de figures représente les 2 modalités de 1 dimension : (8.18a) le *orientation* du plan contenant les cibles, (8.18b) le *cadre de référence* des coordonnées des cibles, (8.18c) le *ancrage* des cibles et leur position (8.18d) la *taille* des cible qui peut être absolue ou relative, et (8.18e) l'*explicité* de l'association cible/commande.

Étude utilisateur 1 : Préférence subjective des stratégies d'affichage des cibles

Nous avons mené une étude auprès des utilisateurs afin d'évaluer les différentes stratégies d'affichage en fonction de leur intuitivité subjective et de leur cohérence par rapport à la tâche.

Résultats et discussion

Le tableau 7.1 représente les résultats du jugement majoritaire pour chacune des 32 stratégies d'affichage. Grâce à cette méthode de classement, nous avons pu regrouper les vidéos en 4 groupes selon les valeurs de 2 dimensions ("Orientation" et "Cadre de référence") (voir Figure 7.1). Les groupes de vidéos avec une image de référence ex-centrée et avec une orientation transversale ont tous une évaluation de majorité "bonne" (vidéos de 25 à 32). Le groupe avec orientation frontale et cadre de référence exocentré a un jugement majoritairement " satisfaisant " alors que les deux autres groupes ont une évaluation majoritairement " faible " ou " acceptable ".

Ces résultats suggèrent que les participants préfèrent une orientation transversale ainsi qu'un référentiel de coordonnées des cibles fixé au robot (exocentré).

Étude d'utilisateurs 2 : Influence de la stratégie d'affichage des cibles sur le rendement de la BCI

Pour évaluer quantitativement l'influence de la stratégie d'affichage des cibles sur la performance de la classification, nous avons réalisé une étude auprès de 24 utilisateurs. Nous avons sélectionné les stratégies DS0, DS15, DS22 et DS29 correspondant à :

- DS0 : Stratégie d'affichage frontale, egocentrée, ancrée sur l'utilisateur, de taille adaptative et implicite (voir Figure 7.3a)
- DS15: Stratégie d'affichage frontale, exocentrée, ancrée sur le robot, de taille absolue et explicite (voir Figure 7.3b)
- DS22: Stratégie d'affichage transversale, egocentrée, ancrée sur le robot, de taille absolue et implicite (voir Figure 7.3c)
- DS29: Stratégie d'affichage transversale, exocentrée, ancrée sur le robot, de taille adaptative et explicite (voir Figure 7.3d)

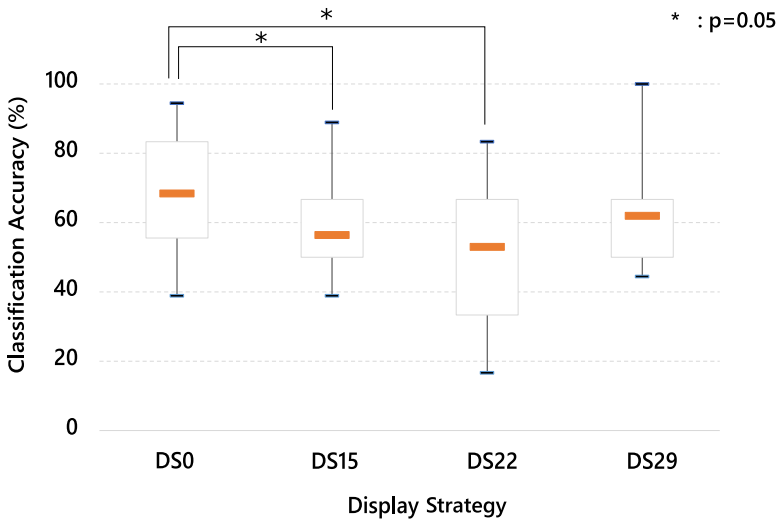


Figure 8.19: Résultats de l'étude utilisateur 2 : Boxplots représentant la précision de la classification, en %, (N=21) en fonction de la stratégie d'affichage.

DS0 est considéré ici comme la condition de contrôle, puisque cette stratégie d'affichage est similaire à une phase d'entraînement standard à la SSVEP et connue pour maximiser la réponse à la SSVEP [Ng et al., 2012].

Résultats et discussion

Le principal résultat de cette expérience est que la stratégie d'affichage des cibles préférée par les participants (DS29) dans l'étude utilisateur 3, donne également de très bons résultats en termes de performance BCI, étant similaire à la stratégie de contrôle et d'affichage optimal DS0 (voir Figure 8.19).

Dans l'ensemble, nos résultats confirment la recommandation d'afficher les cibles dans le même plan que celui du robot (horizontalement) et avec un cadre de référence exo-centré, car cette stratégie est à la fois préférée par les utilisateurs et donne le même niveau de performance que les stratégies d'affichage plus classiques.

Prototype Final

Afin d'illustrer ces résultats dans un cadre pratique, nous avons réalisé un prototype permettant le contrôle d'un robot mobile en utilisant la RA et une ICO.

Figure 8.20: Illustration de notre prototype final en service. (A gauche) Vue générale de l'installation avec l'utilisateur équipé d'un EEG, assis et face au robot mobile réel. (Droite) Vue à la première personne, telle que vue depuis l'HoloLens en utilisant la stratégie d'affichage DS0. La ligne en pointillé représente la trajectoire que le robot a parcourue pendant les sessions de test.

Le robot a été commandé en utilisant les trois mêmes commandes directionnelles : (1) vers l'avant pour le faire avancer à une vitesse de 0,2 m/s, (2) vers la droite pour lui faire effectuer une rotation vers la droite de 15deg/s et (3) vers la gauche pour la rotation vers la gauche à 15 deg/s. Les commandes consistaient en trois cibles SSVEP. Ces cibles étaient activées et affichées lorsque le repère de référence placé sur le robot se trouvait dans le champ de vision de l'utilisateur. La mise au point sur l'une des cibles déclenchait la commande qui lui était associée (voir Figure 8.20).

Conclusion

Dans ce chapitre, nous avons proposé un espace de conception étendu pour l'intégration des cibles dans les applications basées sur la RA-SSVEP. Une premirèe étude a notamment montré que les participants préféraient globalement que le référentiel de la disposition des cibles soit fixé sur le robot (plutôt que sur l'utilisateur).

Une deuxième étude a montré que la disposition des cibles avait un effet sur le rendement de l'ICO. Et enfin, ce travail a illustré la possibilité de développer un prototype complet et opérationnel pour la commande d'un robot mobile en RA à partir d'une ICO.

COMBINER ICO ET AR POUR LA DOMOTIQUE

Introduction

Dans ce chapitre, notre objectif est de proposer des lignes directrices générales pour le développement de systèmes de domotique intelligente AR-BCI, de manière à ce qu'ils soient opérationnels et qu'il soit possible de les mettre à niveau avec les améliorations technologiques et scientifiques futures. Les lignes directrices présentées ci-après sont introduites selon une approche descendante. Nous commençons par examiner les spécifications générales, fonctionnelles et techniques : quelles sont les exigences techniques et les caractéristiques que le système devrait offrir ? Ensuite, nous présentons une architecture de système générique, décrivant ses différentes composantes et leur interaction. Enfin, nous illustrons la mise en œuvre de cette architecture en décrivant un prototype opérationnel spécifique développé avec la société *Orange*, dans un contexte industriel.

Architecture générique

Afin de répondre aux objectifs, nous proposons une architecture générique basée sur 4 composants : (1) La BCI ; (2) le système de RA ; (3) la plate-forme de domotique ; et (4) le Middleware (voir Figure 8.21)

L'ICO

Selon le paradigme BCI souhaité à mettre en œuvre, ce composant met en œuvre le même pipeline décrit dans l'introduction (voir Figure 2) en termes de technique de mesure (le matériel), de traitement du signal et de méthodes de classification.

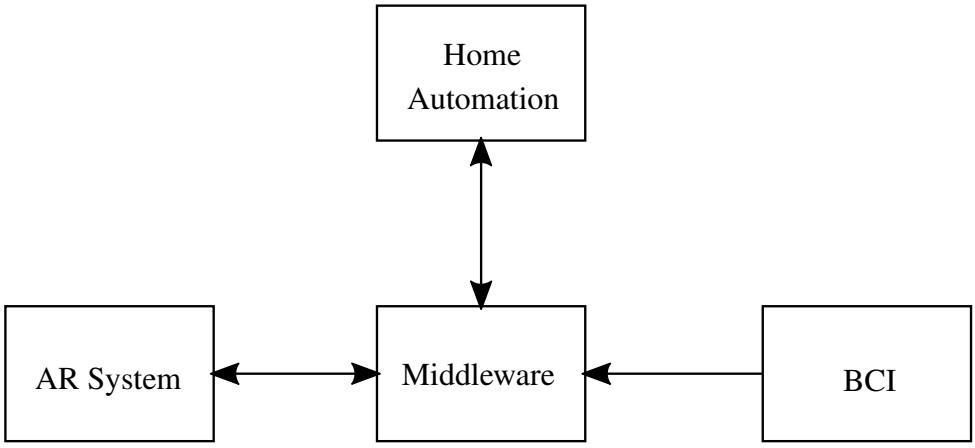


Figure 8.21: Illustration des composants d'une architecture générique d'un système domotique utilisant la BCI et la RA. La communication entre les éléments se fait uniquement par le biais du Middleware.

Le système de RA

Le rôle de cette composante est double. Tout d'abord, il fournit aux utilisateurs des informations sur l'état du système, ainsi que les commandes possibles à envoyer, en fonction de leur contexte (localisation, objets dans le champ de vision par exemple).

Deuxièmement, en plus de la rétroaction et de l'information, cette composante est également chargée de fournir la stimulation visuelle (au moins dans le cas des paradigmes visuels réactifs de l'ICO) pour moduler l'état mental.

La plate-forme domotique

Ce composant est l'interface entre le système et les appareils connectés. Son rôle est de traduire les commandes obtenues à partir de l'interprétation de l'état mental de l'utilisateur déterminé par le BCI et du contexte donné par le système de RA. Si les objets intelligents disponibles ne partagent pas les mêmes protocoles, c'est le rôle de la plate-forme d'automatisation domestique d'adapter les commandes en conséquence.

L'intergiciel

Toutes les composantes décrites ci-dessus ont la particularité d'être indépendantes les unes des autres. Pour pouvoir interagir, l'architecture du système doit prévoir une composante de communication. Essentiellement, le rôle du middleware est de déterminer, à partir du contexte

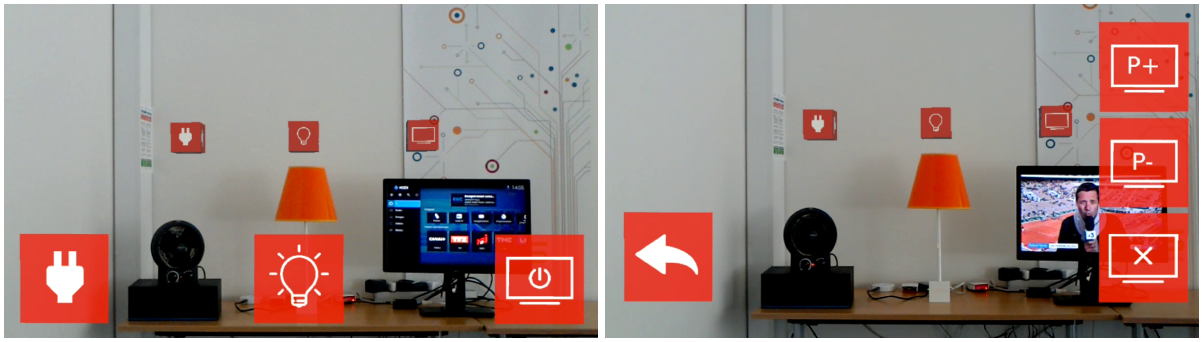


Figure 8.22: Illustration de l'interface RA implémentée. La vue par défaut du système (Gauche) représente les différents objets dans le champ de vision, avec les icônes clignotantes associées. Le ventilateur et la lumière peuvent être allumés ou éteints par une seule commande. L'interaction avec la télévision s'est faite par le biais d'un menu hiérarchique. Après avoir sélectionné le téléviseur, les commandes possibles à émettre apparaissaient sur l'interface (Droite).

des utilisateurs et de leur état mental, la commande correspondante à porter à la plateforme domotique.

C'est l'intergiciel qui intègre le " renseignement " du système puisqu'il est chargé d'agréger les informations provenant des autres composantes, afin de déterminer la commande appropriée. Il est possible d'assimiler son fonctionnement à celui d'un automate à états finis. A partir de l'état de repos initial, les messages provenant des composants changent son état et déclenchent l'envoi d'une commande.

Prototype de domotique

Le prototype a été construit dans le cadre d'un projet de partenariat industriel, avec la société Orange, l'un des principaux fournisseurs d'accès Internet en France. L'objectif était de construire un système domotique opérationnel basé sur RA-BCI, qui serait intégré à leur plate-forme domotique commerciale *Home'in*.

Ce projet a été mené en collaboration avec *Inria Tech*, un service dédié au transfert de technologie des centres de recherche vers l'industrie. Dans ce cadre, notre rôle particulier a été de superviser la conception du système, le développement du BCI ainsi que de fournir les lignes directrices générales. Le développement et l'implémentation des composants logiciels, notamment le middleware, ont été réalisés par des ingénieurs logiciels dédiés.

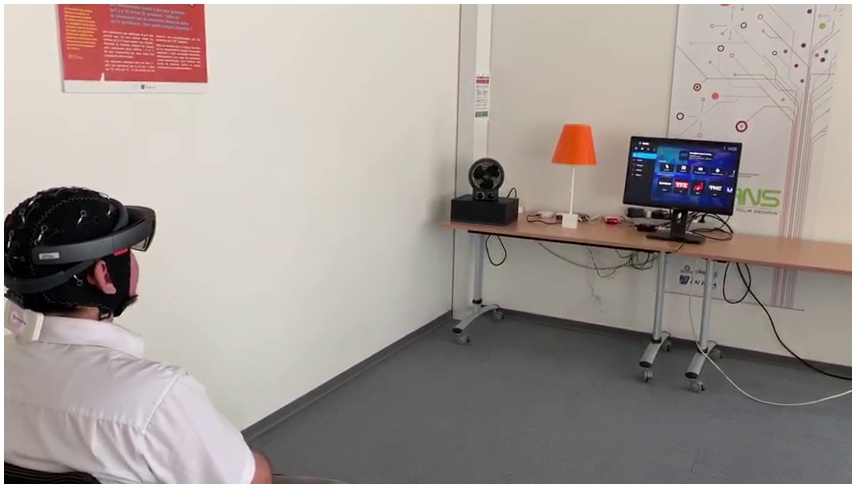


Figure 8.23: Illustration du prototype final. L'utilisateur peut interagir en ligne avec les trois objets devant lui.

Tests informels

Les résultats de nos tests informels sur 2 participants effectués en ligne (voir Figure 8.23), ont confirmé que le système était fonctionnel et a permis de faire fonctionner un scénario d'environ 10 commandes. Un participant a même réalisé le scénario avec une précision parfaite.

Conclusion

Dans ce chapitre, nous avons présenté des lignes directrices qui appuient la mise en œuvre du prototype que nous avons conçu afin d'appliquer nos hypothèses sur la combinaison de l'ICB et de la RA dans un cas d'utilisation réel. Nous avons choisi la domotique, car elle constitue l'un des domaines qui offre les scénarios d'application les plus hétérogènes, compte tenu de la diversité des appareils électroménagers existants.

Nous avons fourni les spécifications générales auxquelles un système combinant la RA et la BCI devrait répondre, et proposé une architecture générique et modulaire, pour faciliter la conception de prototypes opérationnels utilisant la BCI et la RA pour la domotique. Ces lignes directrices et cette architecture ont été regroupées en un système opérationnel, qui a été intégré à une plate-forme domotique disponible dans le commerce.

CONCLUSION

L'objectif global de ce travail était d'étudier l'utilisation des ICO dans la RA. En particulier, les questions sous-jacentes en termes de paradigmes des ICO, de nouveaux types d'interfaces et d'utilisabilité du système. Dans ce but, nous avons identifié 5 objectifs principaux :

1. Introduire une nouvelle perspective des ICO.
2. Étudier la combinaison des ICO et de la RA en fonction des paradigmes des ICO.
3. Améliorer les ICO elles-mêmes.
4. Concevoir de nouvelles interfaces pour l'utilisation d'un ICO dans la RA.
5. Développer de nouveaux usages des ICO et de la RA.

Afin d'atteindre ces objectifs, nous avons proposé dans **Partie II**, une nouvelle définition des ICO qui met en évidence la possibilité de transformer l'activité cérébrale en entrées de processus informatiques (dans **Chapitre 3**).

Dans **Partie III**, nous avons étudié l'utilisation de deux paradigmes d'ICO dans la RA. Nous avons d'abord évalué la possibilité d'utiliser le paradigme de la SSVEP dans la RA. Nous avons étudié l'effet de l'utilisation d'un OST-HMD sur l'élicitation et la détection des réponses SSVEP.

Ensuite, nous avons étudié la possibilité d'obtenir et de détecter les ErrP dans la RA. En particulier, nous avons montré que les erreurs de suivi en réalité virtuelle/accentuée provoquent effectivement de fortes ErrP et qu'il était possible de détecter avec précision la présence de ces ErrP sur la base d'un seul essai.

Dans **Partie IV**, nous avons approfondi l'étude de l'utilisation du paradigme de la SSVEP dans la RA en suivant une approche consistant à (1) améliorer la performance de la ICO basée sur la SSVEP elle-même, (2) améliorer l'intuitivité de l'interface utilisateur de la SSVEP dans la RA, et (3) améliorer l'application en utilisant une ICO basée sur la SSVEP dans la RA.

Notre approche a fourni des indications sur la façon de concevoir des systèmes de RA basés sur des ICO, tout en soulevant de nombreux défis à relever avant que de tels systèmes puissent être utilisés à grande échelle.

Travaux et perspectives d'avenir

Le travail effectué au cours de cette thèse a laissé des possibilités d'amélioration ainsi que quelques questions ouvertes.

Dans la continuité de ce travail, nous pouvons proposer plusieurs pistes d'amélioration à explorer. En particulier, il serait intéressant d'étudier la compatibilité d'autres paradigmes de BCI avec les environnements de RA.

Par exemple, comme il a été démontré que les mouvements de la tête nuisaient considérablement à la précision de la classification SSVEP, les travaux futurs pourraient se concentrer sur l'élaboration de méthodes de filtrage du signal plus efficaces et de procédures d'élimination des artefacts afin de permettre des mouvements plus larges et plus rapides tout en utilisant les ICO. Cela permettrait, par exemple, de sélectionner des cibles se déplaçant plus rapidement. De plus, afin d'accroître la convivialité du système, il serait intéressant d'étudier la possibilité d'utiliser la SSVEP pendant la marche et d'utiliser un système RA-ICO.

Nous étendrions également l'espace de conception proposé pour les interfaces SSVEP-RA à d'autres tâches en plus de la direction de robots mobiles. Par exemple, la question de l'intégration de cibles pour contrôler des objets mobiles 3D (p. ex. des drones volants) ou des dispositifs multifonctions serait intéressante et importante à aborder, afin de rendre l'interaction basée sur la RA-ICO utile dans d'autres cas d'application.

Au jour où ces lignes sont écrites, nous sommes encore très loin d'une machine d'interaction de type Matrix ou basée sur la "pensée" à proprement parler. Malgré des progrès évidents et significatifs dans l'analyse et le décodage de l'activité cérébrale pour déduire les commandes de l'utilisateur, il est encore difficile d'interagir de façon transparente et intuitive avec le cerveau. Et la question demeure : Y parviendrons-nous ? "*Un jour ces réponses pourrons nous sembler aussi évidentes que la terre en orbite autour du soleil - ou peut-être aussi ridicules qu'une tour de tortue. Seul le temps (quel qu'il soit) nous le dira*" [Hawking, 2009].

INSA de RENNES
Service des Formations

AVIS DU JURY SUR LA REPRODUCTION DE LA THESE SOUTENUE

Titre de la thèse:
Design and Study of Interactive Systems based on Brain-Computer Interfaces and Augmented Reality

Nom Prénom de l'auteur : SI MOHAMMED HAKIM

Membres du jury :
- Monsieur BILLINGHURST Mark
- Monsieur SCHERER Reinhold
- Madame SCHULTZ Tanja
- Monsieur LECUYER Anatole
- Monsieur ARNALDI BRUNO
- Monsieur CASIEZ Géry
- Monsieur CABESTAING François
- Monsieur ARGELAGUET Ferran

Président du jury : Bruno ARNALDI

Date de la soutenance : 03 Décembre 2019

Reproduction de la these soutenue

☒ Thèse pouvant être reproduite en l'état
☐ Thèse pouvant être reproduite après corrections suggérées

Fait à Rennes, le 03 Décembre 2019

Signature du président de jury

Le Directeur,

M'hamed DRISSI

UNIVERSITE BRETAGNE LOIRE / MATHSTIC

INSA RENNES

Titre : Conception et Etude de Systèmes Interactifs basées sur les Interfaces Cerveau-Ordinateur et la Réalité Augmentée

Mot clés : Interfaces Cerveau-Ordinateur (ICO), Réalité Augmentée (RA), Systèmes Interactifs, Potentiels Evoqués Visuels Stationnaires (SSVEP), Potentiels d'Erreurs (ErrP)

Résumé : Les Interfaces Cerveau Ordinateur (ICO) permettant l'interaction à partir de l'activité cérébrale. La Réalité Augmentée (RA) elle, permet d'intégrer des éléments virtuels dans un environnement réel. Dans cette thèse, nous avons cherché à concevoir des systèmes interactifs exploitant des ICO dans des environnements RA, afin de proposer de nouveaux moyens d'intéragir avec des éléments réels et virtuels. Dans la première partie de cette thèse, nous avons étudié la possibilité d'extraire différents signaux cérébraux dans un contexte de RA. Nous avons ainsi montré qu'il était possible d'exploiter les Potentiels Evoqués Visuels Stationnaires (SSVEP) en RA. Puis, nous avons montré la possibilité d'extraire des Potentiels d'Erreur des signaux cérébraux, lorsqu'un utilisateur est soumis à des types d'erreurs fréquents en RA. Dans la seconde partie, nous avons approfondi nos recherches sur l'utilisation des SSVEP pour l'interaction en RA. Nous avons notamment proposé *HCCA*, un nouvel algorithme permettant la reconnaissance asynchrones de réponses SSVEP. Nous avons ensuite étudié la conception d'interfaces de RA, pour des systèmes interactifs, intuitifs performants. Enfin nous avons illustrer nos résultats à travers le développement d'un système de domotique utilisant les SSVEP et la RA, qui s'intégre à une plateforme de maison intelligente industrielle.

Title: Design and Study of Interactive Systems based on Brain-Computer Interfaces and Augmented Reality

Keywords: Brain-Computer Interfaces (BCI), Augmented Reality (RA), Interactive Systems, Steady-State Visual Evoked Potentials (SSVEP), Error-Related Potentials (ErrP)

Abstract: Brain-Computer Interfaces (BCI) enable interaction directly from brain activity. Augmented Reality (AR) on the other hand, enables the integration of virtual elements in the real world. In this thesis, we aimed at designing interactive systems associating BCIs and AR, to offer new means of hands-free interaction with real and virtual elements. In the first part, we have studied the possibility to extract different BCI paradigms in AR. We have shown that it was possible to use Steady-State Visual Evoked Potentials (SSVEP) in AR. Then, we have studied the possibility to extract Error-Related Potentials (ErrPs) in AR, showing that ErrPs were elicited in users facing errors, often occurring in AR. In the second part, we have deepened our research in the use of SSVEP for direct interaction in AR. We have proposed *HCCA*, a new algorithm for self-paced detection of SSVEP responses. Then, we have studied the design of AR interfaces, for the development of intuitive and efficient interactive systems. Lastly, we have illustrated our results, through the development of a smart-home system combining SSVEP and AR, which integrates in a commercially available smart-home system.