



UNIVERSITÀ DEGLI STUDI DI MILANO

DIPARTIMENTO DI INFORMATICA

SCUOLA DI DOTTORATO IN INFORMATICA
CORSO DI DOTTORATO IN INFORMATICA XXVI CICLO

TESI DI DOTTORATO DI RICERCA

Towards Steady-State Visually Evoked Potentials Brain-Computer Interfaces for Virtual Reality environments explicit and implicit interaction

SSD: INF/01

AA: 2012/2013

A DISSERTATION PRESENTED

BY

ENRICO CALORE

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN THE SUBJECT OF

COMPUTER SCIENCE


GRADUATE SCHOOL

COORDINATOR:

PROF. ERNESTO DAMIANI

THESIS ADVISOR:

PROF. DANIELE MARINI

 2014 - ENRICO CALORE

THIS WORK IS LICENSED UNDER THE CREATIVE COMMONS ATTRIBUTION NON-COMMERCIAL SHARE-ALIKE 4.0 INTERNATIONAL LICENSE ¹.

Attribution: You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

Non-Commercial: You may not use the material for commercial purposes.

Share-Alike: If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.

<http://creativecommons.org/licenses/by-nc-sa/4.0/legalcode>

¹This license do not apply to figures owned by third parties which are identifiable by the “*courtesy of*” or “*taken from*” clause in the respective captions.

*Towards Steady-State Visually Evoked Potentials
Brain-Computer Interfaces for Virtual Reality environments
explicit and implicit interaction*

ABSTRACT

In the last two decades, Brain-Computer Interfaces (BCIs) have been investigated mainly for the purpose of implementing assistive technologies able to provide new channels for communication and control for people with severe disabilities. Nevertheless, more recently, thanks to technical and scientific advances in the different research fields involved, BCIs are gaining greater attention also for their adoption by healthy users, as new interaction devices.

This thesis is dedicated to the latter goal and in particular will deal with BCIs based on the Steady State Visual Evoked Potential (SSVEP), which in previous works demonstrated to be one of the most flexible and reliable approaches. SSVEP based BCIs could find applications in different contexts, but one which is particularly interesting for healthy users, is their adoption as new interaction devices for Virtual Reality (VR) environments and Computer Games.

Although being investigated since several years, BCIs still poses several limitations in terms of speed, reliability and usability with respect to ordinary interaction devices. Despite of this, they may provide additional, more direct and intuitive, explicit interaction modalities, as well as implicit interaction modalities otherwise impossible with ordinary devices.

This thesis, after a comprehensive review of the different research fields being the basis of a BCI exploiting the SSVEP modality, present a state-of-the-art open source implementation using a mix of pre-existing and custom software tools. The

proposed implementation, mainly aimed to the interaction with VR environments and Computer Games, has then been used to perform several experiments which are hereby described as well.

Initially performed experiments aim to stress the validity of the provided implementation, as well as to show its usability with a commodity bio-signal acquisition device, orders of magnitude less expensive than commonly used ones, representing a step forward in the direction of practical BCIs for end users applications.

The proposed implementation, thanks to its flexibility, is used also to perform novel experiments aimed to investigate the exploitation of stereoscopic displays to overcome a known limitation of ordinary displays in the context of SSVEP based BCIs.

Eventually, novel experiments are presented investigating the use of the SSVEP modality to provide also implicit interaction. In this context, a first proof of concept *Passive* BCI based on the SSVEP response is presented and demonstrated to provide information exploitable for prospective applications.

Contents

1	INTRODUCTION	1
2	BRAIN-COMPUTER INTERFACES	5
2.1	Neurophysiological background	8
2.2	EEG signals acquisition	27
2.3	BCIs modalities	38
2.4	BCIs categories	47
3	BCIs IN VIRTUAL REALITY AND COMPUTER GAMES	53
3.1	General architecture	54
3.2	Active and Reactive BCIs applications	57
3.3	Passive BCIs and Human-Machine Systems	64
4	STEADY STATE VISUAL EVOKED POTENTIALS	77
4.1	Stimuli presentation	81
4.2	Response characterization	94
4.3	Signal analysis	107
4.4	Photosensitive epilepsy	124
5	HARDWARE AND SOFTWARE TOOLS	127
5.1	Acquisition devices	128
5.2	Stimuli presentation devices	131
5.3	The OpenVibe Software	135
5.4	Stimuli presentation software development	145
5.5	A complete SSVEP based BCI system	157

5.6	Conclusion	161
6	PERFORMED EXPERIMENTS	163
6.1	SSVEP BCI using the MindSet	164
6.2	SSVEP elicitation by means of stereoscopic displays	174
6.3	Towards SSVEP based <i>Passive</i> BCIs	184
7	CONCLUSIONS	209
	REFERENCES	215

1

Introduction

Brain-Computer Interfaces (BCIs) implement a direct communication pathway between the brain and an external system using hardware bio-sensors able to record the neural activity and software tools able to extract from the recorded signals information regarding particular brain states. To each detectable brain state could be associated a command for a generic system which could implement communication or control functions directly controlled by the neural brain activity of the user.

At the moment different research groups working on BCI research exist ¹ and in Europe most of them are focused on non-invasive brain recordings, which are performed mainly using electroencephalographic devices. This kind of recording technique is the most suited to allow practical BCIs for healthy users, which could be operated also in an out-of-the-lab environment using devices available at an affordable cost.

BCI research involves a wide range of different independent research fields as

¹For example: <http://bci.tugraz.at/> and <http://www.bbci.de/>

neuroscience, neuro-physiology, bio-engineering, mathematics and computer science, thus most important research centers are commonly born melting competences from different university departments. BCI research is even increasingly involving more research fields as new applications are proposed, attracting researchers from the Human-Computer Interaction (HCI) field and involving theories and results from the psychology and human factors research areas.

The high level of multidisciplinary, as highlighted also by Dr. Thorsten O. Zander², causes the initiation of BCI research from scratch to be extremely challenging and indeed he has been a strong promoter for the foundation of interdisciplinary research groups in order to bring students of psychology, mathematics, human factors and engineering to cooperate towards the common goal of BCI research.

Since also in my Department the research in this field is being started recently, a preliminary study reviewing the fundamental knowledge at the basis of a generic BCI, spanning different research areas, has been performed in this work. A second study has later been accomplished with the goal to provide a review of existing BCI implementations in the context of Virtual Reality (VR) environments and Computer Games, which are the applications of main interest for this thesis.

According to the results available in the literature, a particular BCI modality has later been selected, as considered the most suited for my goals and has been studied in depth in order to highlight the characteristics that a state-of-the-art implementation should possess. Eventually, an implementation is proposed and some experiments are presented aiming to overcome some of the known limitations and to propose new prospective applications.

In the beginning of Cap. 2 a detailed description of the BCI concept is given, highlighting the different names and definitions used by different research groups working towards the same goal of brain controlled systems. Different names and definitions in the literature to identify similar concepts are explained by the fact that, as already mentioned, BCI research is widely multidisciplinary and different groups in the last decades begun to investigate in the same direction starting from very different research fields. After an initial review of the different BCI defini-

²Dr. Thorsten O. Zander is one of the founder of the interdisciplinary BCI research group Team PhyPA <http://www.phypa.org/> at the Technische Universität Berlin, working mainly on *Passive* BCI research.

tions found in literature, in this chapter are addressed: the neuro-physiological basis justifying the possibility to record signals reflecting the brain activity; how these signals could be recorded, focusing mainly on the electroencephalographic technique; and eventually the most used BCI modalities within their different categorization.

Moving in the specific application context of interest, in Cap. 3 are reviewed some BCI implementations, spanning over different modalities and categories, aimed to the interaction with VR environments and Computer Games, highlighting the wide range of possible applications. In this chapter are initially reviewed applications aimed to explicit interaction, while after an introduction to the psychological concept of *flow*, recently implemented and new envisioned applications aimed to implicit interaction are presented as well.

On the other side, moving on a more specific BCI modality, in Cap. 4 is presented an in-depth review of the available literature regarding the Steady State Evoked Potentials (SSVEP). In particular, are included only researches which could be of interest for BCI applications, spanning from the fields of vision research that initiated the SSVEP investigation, to the field of bio-signal processing, reviewing the state-of-the-art algorithms used for their detection. BCI implementations exploiting this modality are reviewed as well, describing the different approaches followed in previous works, trying to identify the possible pitfalls. Along with this review, implications for practical BCI applications will be discussed as well in the attempt to define a state-of-the-art multidisciplinary basis to be used for a SSVEP based BCI practical implementation, able to exploit the pre-existing knowledge in the different fields.

Eventually, Cap. 5 and Cap. 6 are mainly dedicated to my contributions. In particular, Cap. 5 describe the available hardware in the Laboratory and the software tools adopted, chosen within the available ones if considered at the state-of-the-art level, or custom implemented if otherwise.

Finally, Cap. 6 presents experiments aimed to demonstrate the practicability of SSVEP based BCIs in an out-of-the-lab context exploiting commodity hardware and state-of-the-art software. Furthermore, experiments aimed to investigate the possibility of using SSVEP based BCIs as implicit interaction devices are presented in this chapter as well.

2

Brain-Computer Interfaces

A Brain-Computer Interface (BCI), in its historical, general and wide accepted form, is defined as a direct communication pathway between the brain and an external device as a computer or a machine in general [197]. In different research communities the same concept with slightly different meanings has also been called Brain-Machine Interface (BMI) or Brain Neural Computer Interface (BNCI).

The term BMI can often be found in research works where invasive electrodes are used [2, 135], while BCI is the most used term where non invasive recording is performed. On the other side, BNCI is a broader definition introduced by the European Commission comprising systems exploiting signals acquired also from the peripheral neural system [2].

BCIs can be implemented using different kind of tools capable of reading the brain activity of an user and signal processing algorithms capable of extrapolate from the read activity a particular state of the user's brain. To each detectable state could be associated a command for an actuator. A schematic view of a generic BCI

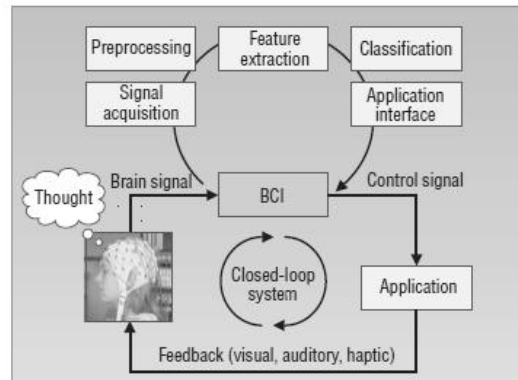


Figure 2.0.1: A generic BCI components scheme. Figure taken from <http://future-bnci.org/>.

system is shown in Fig. 2.0.1.

Most of the research regarding BCIs in the past aimed to provide mobility impaired users with a tool capable of translating a thought or a will into a command for an external device or a prosthetic limb. More recently, thanks to the advancement of the research in the field, both from the software and from the hardware point of view, BCIs use is being investigated also in the field of Human-Computer Interaction (HCI) to provide new communication channels that do not rely on the user's limb movements, also for healthy subjects [185].

As previously mentioned, various definitions exist about the BCI term; some researchers exclude from it the systems which are not real-time or systems that use neural activity detected from peripheral nerves or systems where the user is not willing to instruct a command, but its mental state is passively read for implicit interaction.

A unique definition has been searched for years, but has been very hard to find a consensus between the researchers which started to work on this topic coming from very different research fields. A document has been proposed in 2011 to solve this issue [2], but as stated in it, the discussion is still not over and also in the Asilomar BCI Conference held in June 2013 a questionnaire has been distributed in order to understand what the majority of the researchers in the community identify as a BCI and what according to them should not be named with this term.

As stated in [2], a BCI should meet four criteria:

- **Direct:** The system must rely on direct measures of brain activity. With existing sensor technologies, this means that sensors must be placed in, on, or very near the head, since there are no technologies that can measure brain function from afar. A device is not a BCI if it only acquires information that travels through peripheral nerves or muscles before being detected.
- **Real-time:** Most modern BCIs allow people to send a messages or commands every 2-5 seconds. To account for some BCIs that enable communication having longer selection times, real-time refers to a maximum of a one minute delay between the user's formation of a relevant message or command and resulting feedback. It is possible that BCIs or other communication systems with a longer latency could still be effective near real-time tools, such as a Galvanic Skin Response system. Offline systems are not BCIs.
- **Feedback:** BCIs must present real-time feedback to the user. That is, the system must act on the user's intent so that the user can know whether s/he successfully conveyed the desired message or command. For example, the BCI might present a letter on a monitor, move a wheelchair, affect a virtual environment, or control a robotic device. The "real-time" and "feedback" criteria could be combined into one criterion called "closed-loop".
- **Intentional:** The user must perform some voluntary, intentional, goal directed mental activity each time s/he wishes to convey information. This activity must be for the sole and specific purpose of using the BCI. That is, a system that only acts on brain activity produced as a side effect of performing another task is not a BCI. This feature is not addressed in most articles, and entails further clarification of the phrase "message or command".

In my opinion and also as stated in [2], the 4th criterion is the most controversial and seems that a considerable consensus is growing around the expansion of the BCI definition to encompass also devices based on unconscious and unintentional signals, under the name of *Passive* BCI, first introduced in [39].

Passive BCIs and related terms have been used since the Asilomar BCI Meeting in 2010 and the affective BCI workshop, held at the ACII conference in Memphis

Acquisition method	Measured activity	Temporal resolution	Spatial resolution	Invasive	Portable
EEG	Electrical	~ 0.05 s	~ 10 mm	No	Yes
MEG	Magnetic	~ 0.05 s	~ 5 mm	No	No
ECoG	Electrical	~ 0.003 s	~ 1 mm	Yes	Yes
Intracortical	Electrical	~ 0.003 s	0.005 to 0.5 mm	Yes	Yes
fMRI	Metabolic	~ 1 s	~ 1 mm	No	No
NIRS	Metabolic	~ 1 s	~ 5 mm	No	Yes

Table 2.0.1: Comparison of different methods to acquire signals reflecting the ongoing brain activity of a subject. Data taken from [134].

in October 2011, featured a day of talks and discussion surrounding affective BCIs. Moreover recent articles discussed this new term [199, 201] together with a categorization of different BCIs kinds. Therefore I decided in this work to adopt this categorization, which will be exposed in detail in Sec. 2.4, using the term *Passive* BCIs for BCIs that does not comply with the 4th criterion.

Different kind of devices could be used to record brain activity as Electrocor-ticography (ECoG), Functional Magnetic Resonance (fMRI), Magnetoencepha-lography (MEG), near Infrared Spectroscopy (NIRS), etc., but, despite of this, BCI research aimed to provide practical applications is focused mainly on the use of Electroencephalography (EEG) acquisition devices. As shown in Tab. 2.0.1, EEG acquisition devices could indeed be used in an out-of-the-lab environment, they provide a relatively high time resolution (that is essential for interactive ap-plications), without requiring surgical operations and moreover they are relatively inexpensive. Therefore also in this work, brain activity will be recorded using EEG acquisition devices and further details about this recording method will be given in Sec. 2.2.

2.1 NEUROPHYSIOLOGICAL BACKGROUND

Despite of the kind of acquisition device used in a BCI, the source of the acquired signals has to be given by the activity of neurons in the brain. Therefore in this sec-tion will be given an introduction to the neurophysiological basis related to these

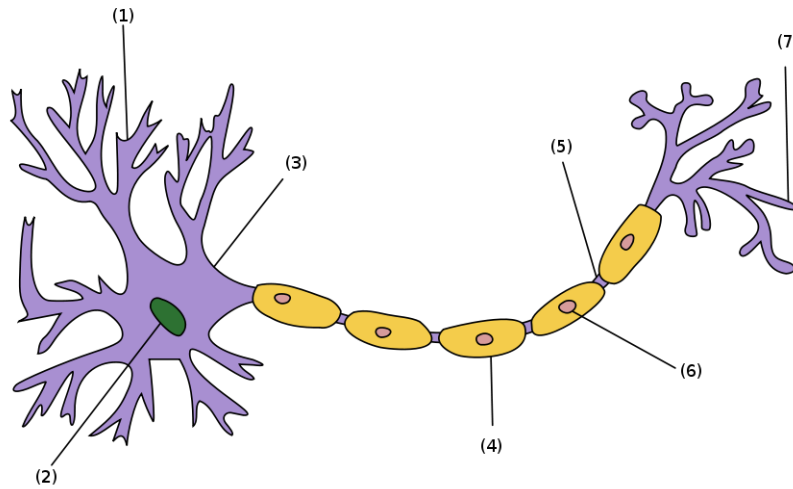


Figure 2.1.1: Structure of a typical neuron. (1) Dendrite; (2) Nucleus; (3) Soma; (4) Axon; (5) Node of Ranvier; (6) Schwann Cell; (7) Axon terminal;

signals generation and how they could be recorded.

The electroencephalogram acquisition will be dealt in particular in Sec. 2.2, since it is the preferred neurological signal used in this work.

2.1.1.1 THE NEURON

The fundamental component of the nervous system, which includes the brain, spinal cord and peripheral ganglia is the neuron, as depicted in Fig. 2.1.1 schematically and sketched by the Nobel prized neuroscientist Santiago Ramón y Cajal, in Fig. 2.1.2.

The human brain contains about 10^{11} neurons and each of them has on average of 7,000 connections to other neurons. It has been estimated that an adult's brain has from 10^{14} to 5×10^{14} connections between neurons.

A neuron is an electrically excitable cell that processes and transmits information through electrical and chemical signals. Neurons are discrete cells, not continuous with other cells and idealizing them as black boxes, we may say that information flows from the dendrites ((1) in Fig. 2.1.1), the inputs of the neurons, to the axon ((4) in Fig. 2.1.1) the output, via the cell body named Soma ((3) in

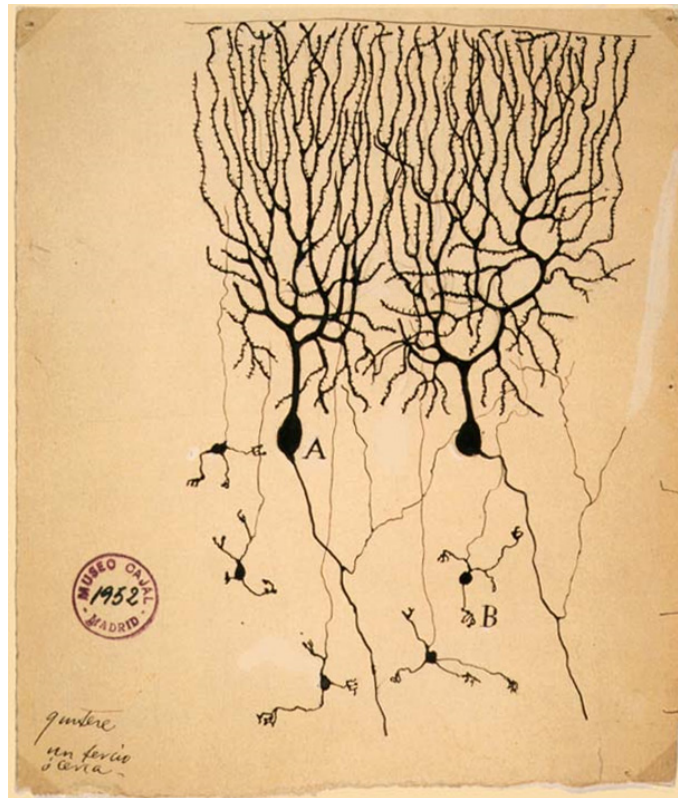


Figure 2.1.2: Drawing of Purkinje cells (A) and granule cells (B) from pigeon cerebellum by Santiago Ramón y Cajal, 1899. Instituto Santiago Ramón y Cajal, Madrid, Spain.

Fig. 2.1.1), that is where a “decision is taken” about the output response given the input signals.

The neurons capability of receiving and transmitting information is given by the fact that neural cells are able to rapidly change the intracellular-versus-extracellular concentrations of several ions such as sodium (Na^+), potassium (K^+) and chloride (Cl^-). The different ions concentrations between the inside of the cell membrane and the outside medium, gives a voltage potential difference that could be changed according to the concentrations ratio and that could be propagated along the axon to reach other neurons’ dendrites.

As stated by the computational neuroscientist Dr. Rajesh P.N. Rao, the neuron could be facetiously defined, from a practical point of view, as a «leaky bag of charged liquid». It is a bag full of charged liquid, since the cell membrane is a

lipid bilayer that is impermeable to aforementioned charged ion species, insulating the inside from the outside of the cell, but it is also “leaky”, since embedded in the membrane there are ionic channels which are a sort of gates allowing some ions to flow in or out, as depicted in Fig. 2.1.3.

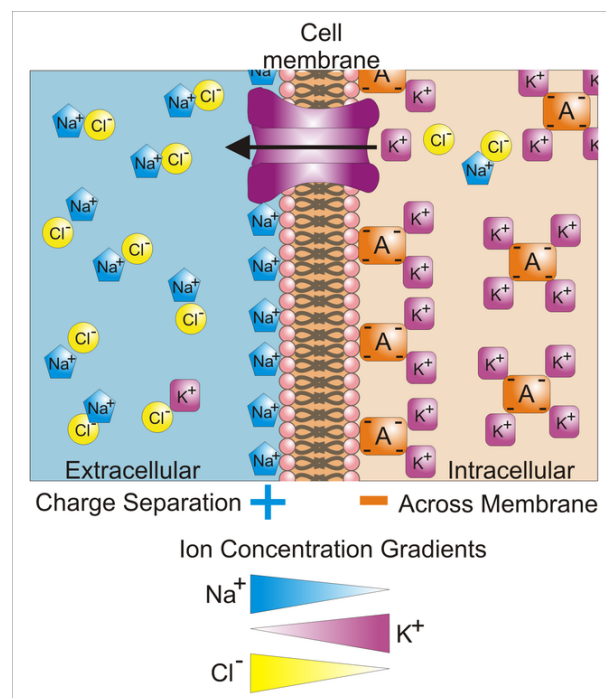


Figure 2.1.3: Schematic representation of the ions concentrations inside and outside the cell membrane of a generic neuron. Figure courtesy of Vojtěch Dostál, cc-by 3.0.

In resting conditions the voltage potential between the inside and the outside of a neuron is about -70 mV, where the inside of the neuron is negatively charged with respect to the outside medium. This difference called resting potential, is due to the operation of particular gates on the cell membrane, called ionic pumps, that actively expel from the cell Na⁺ ions while allowing K⁺ ions in. As a result Na⁺ and Cl⁻ are more concentrated outside of the cell while K⁺ is more concentrated in the inside.

Ionic channels in membranes are proteins that are selective, in the sense that they allow only specific ions to pass through each direction, but in order to allow neurons to receive and transmit information through electrical signals, their behav-

ior has to change in time, in order to alter the membrane potential that otherwise would be constant. Indeed, some ionic channels are able to change their behavior according to the local environment; e.g. three main kind of ionic channels exist:

- *Voltage-gated*: changes of the local membrane potential causes them to open;
- *Chemically-gated*: binding to a chemical causes them to open;
- *Mechanically-gated*: are sensitive to pressure or stretch.

Gated channels allow neuronal signaling and communication between the neurons. Particular junctions between neurons, called synapses, are able to transmit incoming electrical signals from the axon of the transmitter neuron, to the cell body of the receiver neuron, commonly thanks to chemically-gated channels at the junction and in some cases thanks also to voltage-gated channels.

In the case of a chemical synapse, as depicted in Fig. 2.1.4, an electrical impulse coming from the axon of a transmitting neuron, causes the axon terminal to release some molecules called neurotransmitters in the proximity of the membrane of a receiving neurons where chemically-gated channels are. These channels, being sensitive to the neurotransmitter molecules, may open and for example (according to the kind of channel), allow Na^+ ions to enter the receiver neuron's membrane increasing its local membrane potential at that particular membrane location.

A change in the local membrane potential may activate other voltage-gated channels that are nearby on the membrane and according to the kind of voltage-gated channels, this may lead to a stronger increase of the membrane potential, named depolarization, or its decrease named hyperpolarization.

A strong enough depolarization of the cell membrane causes what is called an action potential, or a spike, that translates to the “decision” to transmit a signal over the axon of the receiver neuron (that now became a transmitter itself), to another, or more commonly various, other neurons.

A graph depicting a theoretical and a measured action potential is reported in Fig. 2.1.5. In reference to that graph, the depolarization of the cell membrane leads to Na^+ channels to open and thus a rapid inflow of Na^+ ions to occur; if the depolarization is strong enough, reaching a threshold level of about -55 mV , it causes

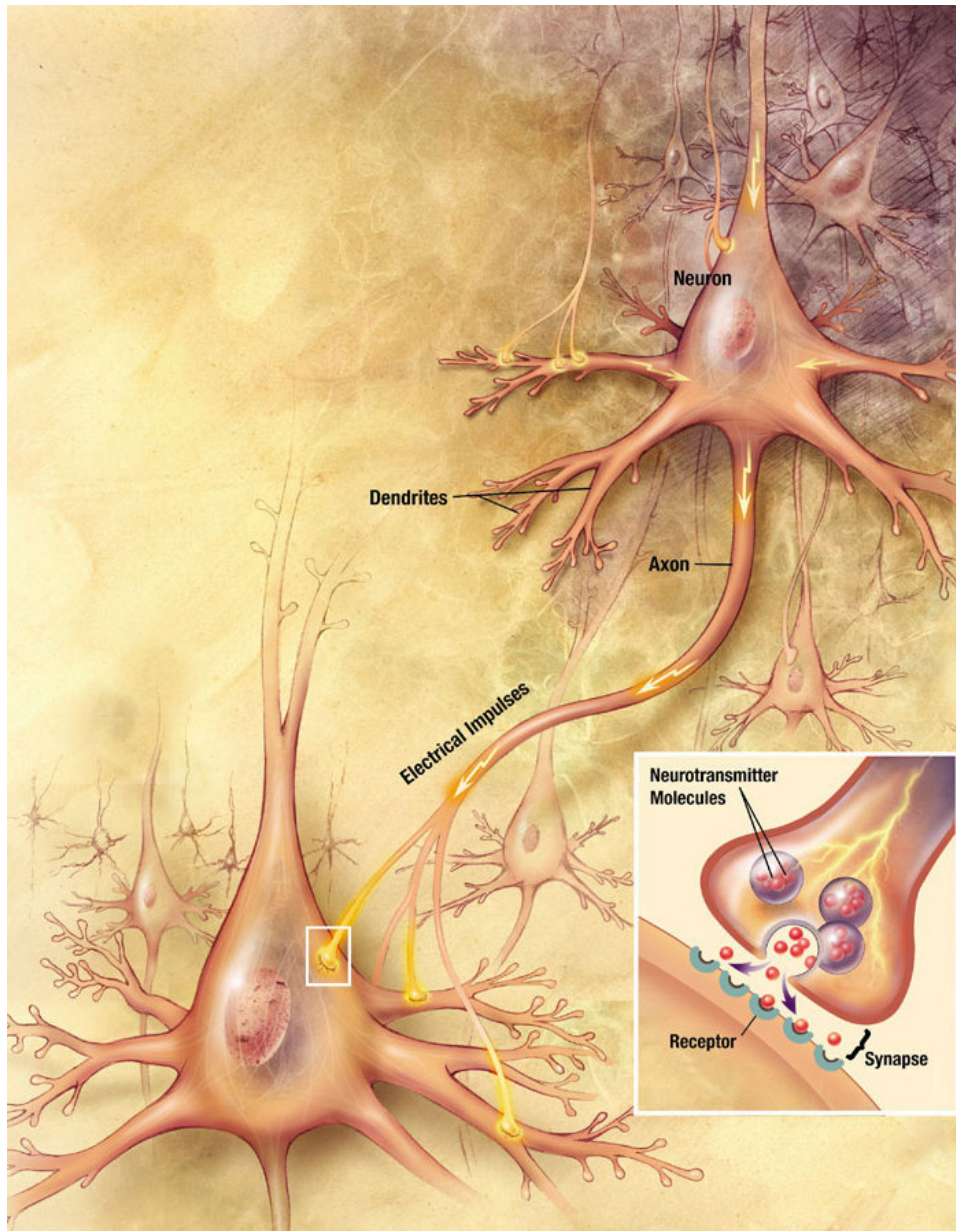


Figure 2.1.4: Graphical representation of the information flow between two generic neurons highlighting the functioning of a chemical synapse. Figure courtesy of US National Institutes of Health, National Institute on Aging.

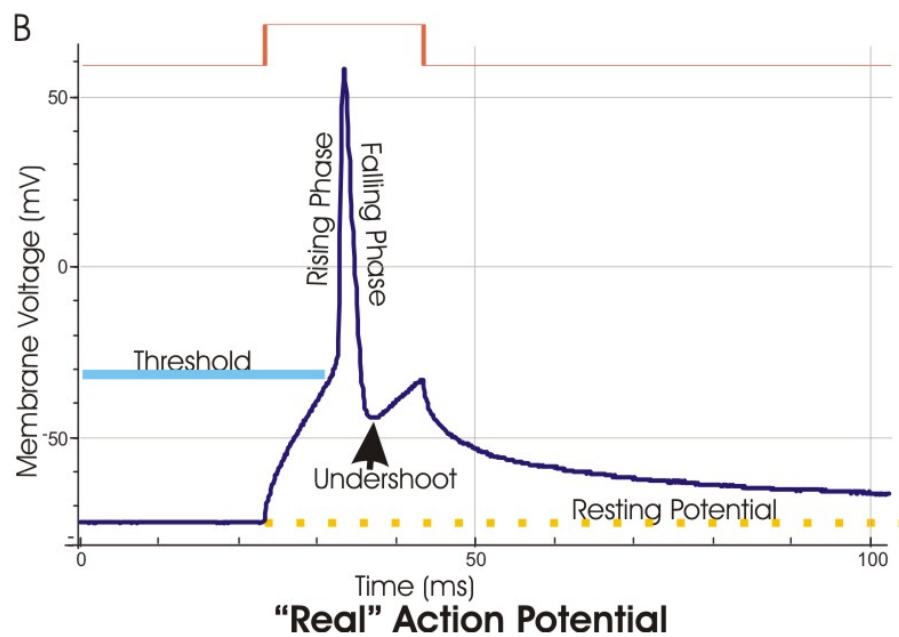
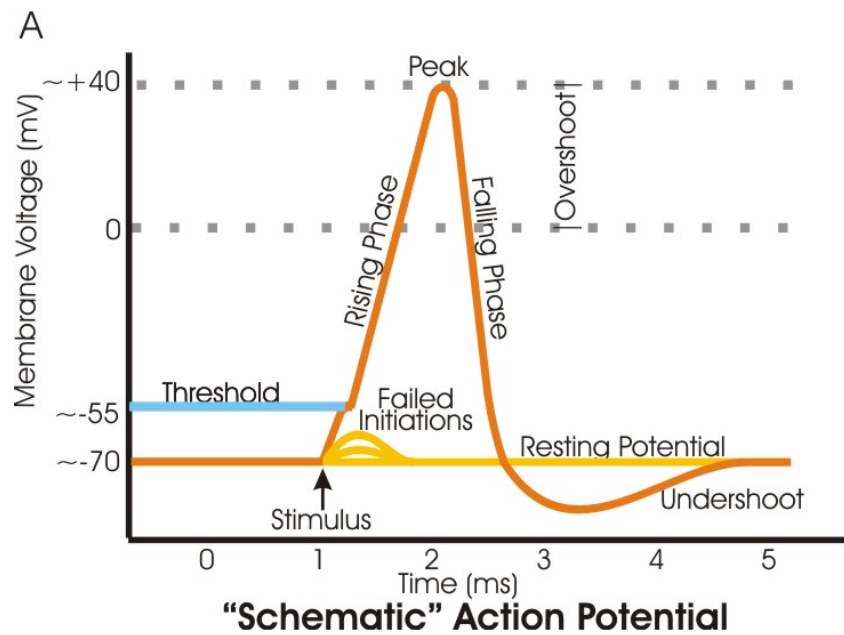


Figure 2.1.5: This figure represent a theoretical and a measured action potential. The membrane potential variation is given in millivolts over the time given in milliseconds. Figure courtesy of <http://en.wikipedia.org/>, cc-by-sa 3.0.

an even stronger depolarization starting a positive feedback loop until a peak of about 40 mV is reached. When the peak is reached Na^+ channels close and about at the same time K^+ channels open, causing an outflow of K^+ ions, rapidly lowering the membrane potential, letting it to reach again its resting potential and leading the K^+ channels to close again.

As soon as a neuron produces an action potential, it is propagated along its axon, till it reaches one or more synapses at the axon terminals ((7) in Fig. 2.1.1) to be transmitted to other neurons. Longer axons (that could be longer than one meter) are often myelinated, meaning that they are covered with a substance called myelin produced by glial cells ((6) in Fig. 2.1.1), in order to insulate them with respect to the outer environment to lower signal loss. Moreover at the nodes of Ranvier ((5) in Fig. 2.1.1) action potentials are received and regenerated, thanks again to sodium and potassium channel gates, implementing an active wire for fast long-range lossless signal propagation.

Being the shape of the action potentials related to the sodium and potassium gates dynamics, it is the same for all the neurons and it does not carry any information. The action potential or spike, can therefore be seen as a single impulse, that may or may not occur at a specific time, but the output of a neuron could indeed be completely described as an impulse train.

2.1.2 NEUROPHYSIOLOGIC BASIS OF EEG

Single neuron action potentials or spikes, cannot be recorded non-invasively, since the extracellular current they generate is too weak to be detected without an implanted electrode positioned near the cell membrane. Despite of this, the synchronous firing of multiple neurons near the surface of the brain, can induce a potential field strong enough to be detected on the skin surface [19].

Hans Berger discovered indeed that is possible to measure electrical activity from the human brain connecting electrodes to the scalp at the beginning of the 20th century. The graphic representation of such electrical activity between two different cerebral locations plotted over time is known as the electroencephalogram or EEG. In Fig. 2.1.6 is reported the first human EEG recording appearing in Berger's first publication on human EEG [8].

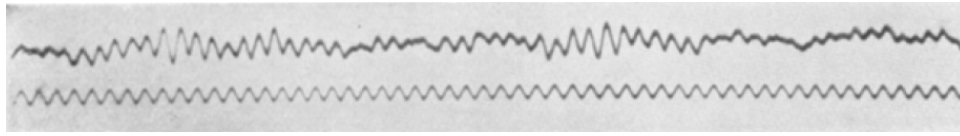


Figure 2.1.6: The first human EEG recording obtained by Hans Berger in 1924. The upper tracing is the EEG, while the lower is a 10Hz timing signal. This image is one of the first EEG recordings, appearing in Berger's first publication on EEG [8], it is a portion of Fig.13.

The most significant sources of EEG potentials are both the excitatory and inhibitory postsynaptic potentials (EPSPs/IPSPs), generated at the end of the axons as shown in Fig. 2.1.7.

The summation of multiple neurons EPSPs and IPSPs generate what is called a Local Field Potential (LFP), measurable by invasive electrode recordings and shown in Fig. 2.1.8, while the summation of larger neuronal populations and thus stronger LFPs, can be detected also over the scalp with non-invasive EEG recordings [19].

The electroencephalogram signal reflects mainly the activity of the cerebral neurons that are closer to the scalp and in particular of the ones perpendicularly oriented with respect to it, as shown in Fig. 2.1.7, since these conditions grant a weaker attenuation. Dendrites which are deeper in the cortex, inside sulci, in midline or other deep structures, or producing currents that are tangential to the skull, have far less contribution to the EEG signal and in some cases they could also elide their contributions by themselves (if of opposite sign/direction).

The EEG signal suffers from several limitations given by the fact that it do not provide an exact representation of the neuronal activity. In fact the EEG signal is an average over the activity of large populations of neurons and moreover it is attenuated and distorted, due to the varying electrical conductivity properties of the tissues between the source and the electrode and due to the different neurons orientations [40, 140].

Although voltage potential differences measured near the neuron membrane, are in the order of some millivolts, surface recorded EEG commonly fluctuates under $\pm 100 \mu\text{V}$ due to conduction attenuation.

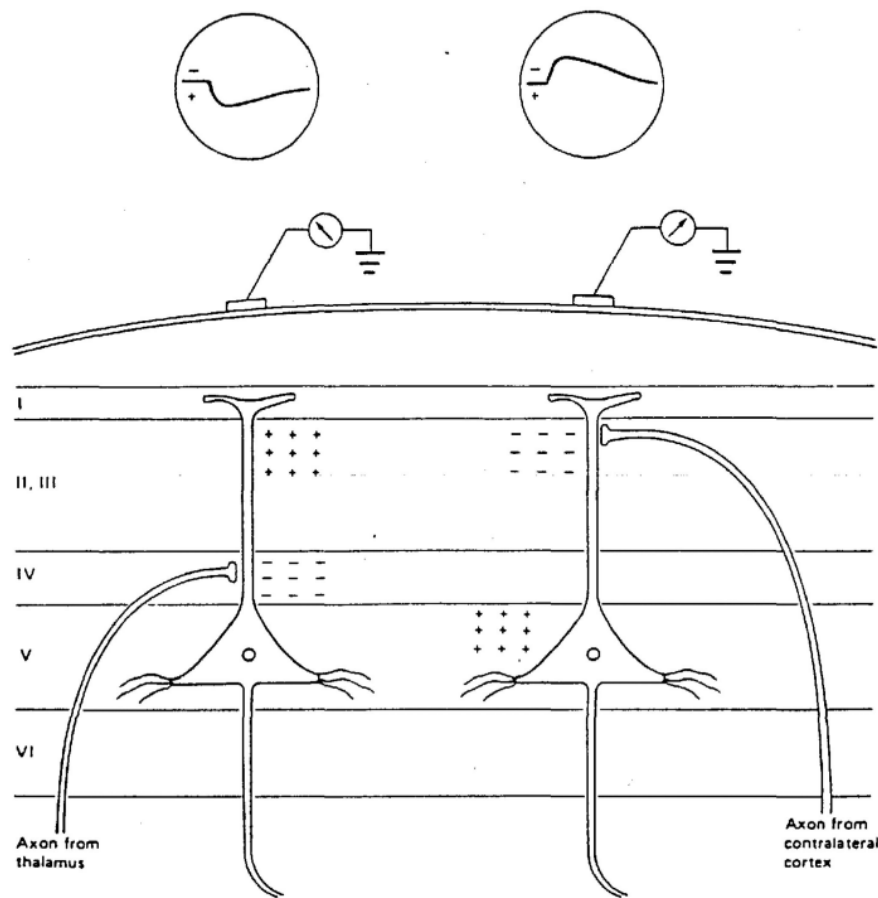


Figure 2.1.7: Generation of extracellular voltage fields from graded synaptic activity. Neurons size is clearly not in scale and they represent the actions of several neurons perpendicularly oriented with respect to the scalp. In the upper part of the figure are depicted two electrodes measuring scalp potentials fluctuations with respect to an ideal ground. Figure taken from [112].

Furthermore, being the electrodes placed on the scalp, the acquired EEG signals are a two-dimensional projection (attenuated and distorted as mentioned before) of the neuronal activity happening in a three-dimensional space (the whole brain). This is commonly referred as the inverse problem, stating that is theoretically impossible to determine the three dimensional signal source localization given only the surface projection (although various algorithms performing source localization exist in order to provide probabilistic estimations).

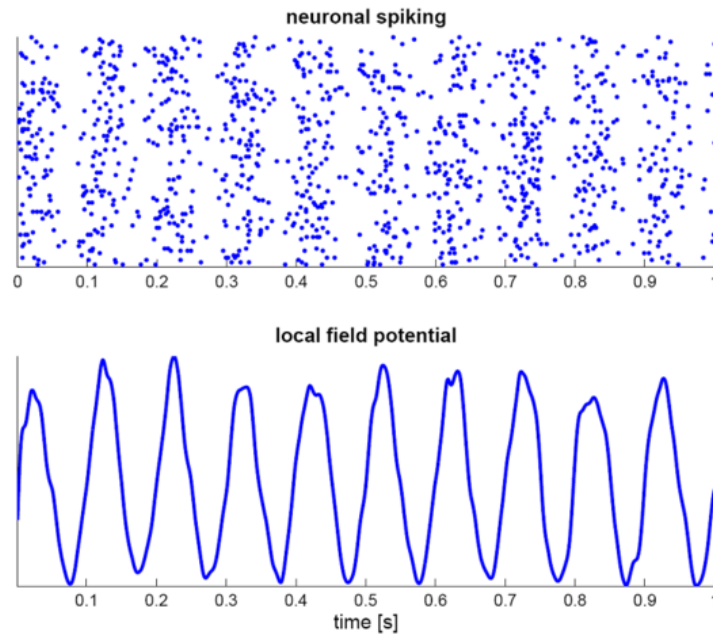


Figure 2.1.8: In the upper side of the figure is reported a simulation of a synchronized neuronal population firing; every row represent a neuron and every point correspond to a neuron's spike with respect to time (on the x-axis). In the lower part of the figure is represented the corresponding Local Field Potential measurable near the neurons population, but also on the scalp for large populations near the surface.

2.1.3 DIFFERENT AREAS OF THE BRAIN

The human cerebral cortex, which is the outer part of the brain and thus the one giving the most contribution to the surface recordable EEG signals, is anatomically divided in four main lobes, as depicted in Fig. 2.1.9.

The temporal and frontal lobe are divided by the Sylvian fissure, while the central sulcus divides the frontal and parietal lobe.

Further divisions of the cerebral cortex based on the cytoarchitectural¹ organization of neurons have been introduced at the beginning of the 20th century as the Brodmann areas depicted in Fig. 2.1.10 which are still commonly referenced and

¹In biology, *cytoarchitecture* refers to the arrangement of cells in a tissue or the molecular construction of a cell. In neuroscience, it refers specifically to the arrangement of neuronal soma in the brain and spinal cord.

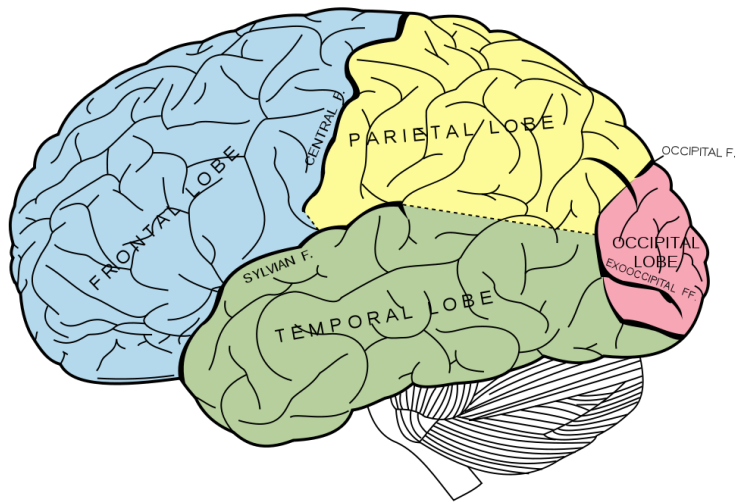


Figure 2.1.9: The four lobes of the cerebral cortex.

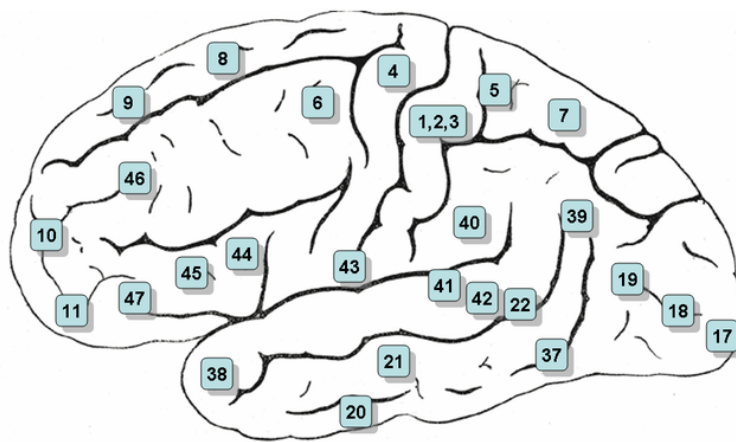


Figure 2.1.10: The Brodmann's division of the cerebral cortex.

used.

Later researches gave evidence of the correlation between the Brodmann cortical areas and local cortical brain functions using neurophysiological and functional imaging methods. For example, Brodmann areas 1, 2 and 3 are the primary somatosensory cortex, area 4 is the primary motor cortex and area 17 is the primary visual cortex.

SOMATOSENSORY CORTEX

In the field of BCI, where different brain activities have to be detected and recognized, it is of great importance to know in which part of the brain each function is happening in order to know where to place surface electrodes (in the case of EEG based BCIs) as will be discussed in Sec. 2.2.1.

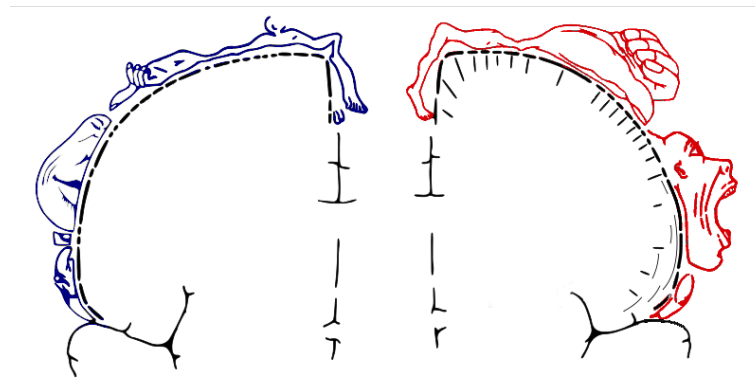


Figure 2.1.11: The “homunculus”, a cortical pictorial representation of the functional divisions of the primary motor cortex (on the right) and the primary somatosensory cortex (on the left). They actually represent two different slices of the brain; both the hemispheres are symmetric and thus exists a left primary motor cortex as well as a right one and the same holds also for the somatosensory cortex.

Finer maps are indeed available and a famous one related to the primary motor and somatosensory cortex is reported in Fig. 2.1.11. In this figure is reported a graphical representation of each part of the human body associated to the underlying part of the cerebral cortex controlling it or receiving stimuli from it. This map is used for example in the case of Motor Imagery (MI) Brain-Computer Interfaces,

which will be addressed in Sec. 2.3.2. The size of the body parts seem distorted since they are not proportional to the part sizes, but to the area of the cortex associated to them.

VISUAL CORTEX

Another part of the cerebral cortex which is of interest for some kind of BCIs using visual stimuli to elicit particular brain reactions (as the ones discussed in Sec. 2.3.1 and Sec. 2.3.3), is the primary visual cortex, often named V_1 , situated in the Brodmann area 17.

The reason why the visual cortex resides in the occipital (rear) part of the brain is explained by the path followed by visual information along the brain, as shown in Fig. 2.1.12.

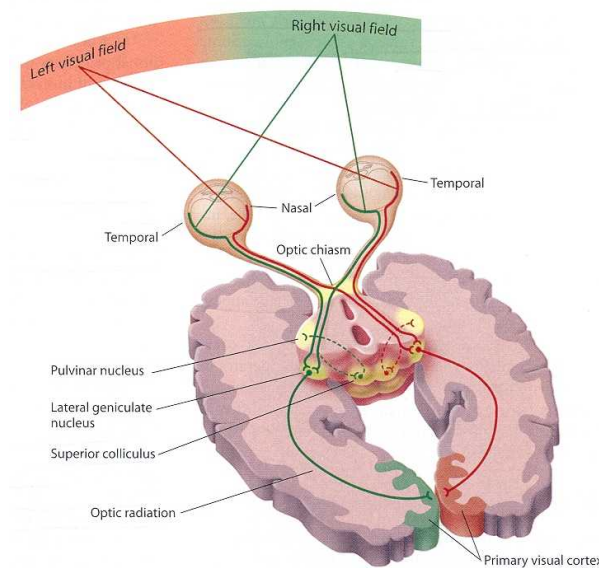


Figure 2.1.12: Sketch of the human Visual Pathway along the brain. Figure courtesy of <https://wiki.ucl.ac.uk/>

The visual cortex is commonly divided into various sub-areas, as the primary visual cortex (also known as striate cortex or V_1) and extrastriate visual cortical areas such as V_2 , V_3 , V_4 , and V_5 (referred also as MT). The extrastriate cortical areas consist of Brodmann area 18 and 19 as shown in Fig. 2.1.13.

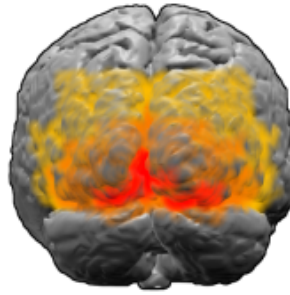


Figure 2.1.13: Brodmann Areas (BA) 17, 18 and 19. BA 17 is shown in red. BA 18 is orange. BA 19 is yellow. This is a rear view of the brain. Much of BA 17 is hidden from view on the medial surface (between the hemispheres), on the ventral bank of the calcarine sulcus. The brain's surface is extracted from structural MRI data (Wellcome Dept. Imaging Neuroscience, UCL, UK). The Brodmann Area data is based on information from the online Talairach demon (an electronic version of Talairach and Tournoux, 1988).

When light hits the retina inside the human eye, it is absorbed by two types of photoreceptors: *rods* and *cones*. The rods are more numerous and sensitive to the light intensity, but does not discriminate between different light wavelengths (i.e. colors). Furthermore, there are very few rods in the center of the eye visual field (named the *fovea*). Concerning the cones, there are three kinds of them which are sensitive to light of different wavelength bands. Approximately 64% are sensitive to green, 32% to red and only 2% to blue light. The red and green cones are mostly concentrated in the fovea, while the blue cones are relatively more sensitive [10].

When photoreceptors are hit by light beams, they may absorb photons, triggering a change in their cell's membrane potential which is detected and processed by retinal neural populations (bipolar cells and amacrine cells) and later transmitted by ganglion cells as action potentials through their axons (i.e. the optic nerve) back in the brain as shown in Fig. 2.1.12 and in Fig. 2.1.14.

Activation from each visual field is sent contralaterally to the lateral geniculate nucleus (LGN) along three different pathways [188].

The M-pathway (named after the *magnocellular* neurons it is connected to) goes through brain areas V₁, V₂, V₅ and STS/PP, it is involved in the detection of coarse and dynamic shapes, motion and depth and represents the “where” part of visual information [10]. It is primarily associated with the rods in the retina.

The P-pathway (after *parvocellular* neurons) is mostly connected to the red and green cones and is involved in the detection of high spatial contrasts, color information (specifically red and green) and details. Moving through the V₁, V₃, V₄ and IT areas of the brain, it is slower than the M-pathway and represents the “what” part of visual information [10, 80].

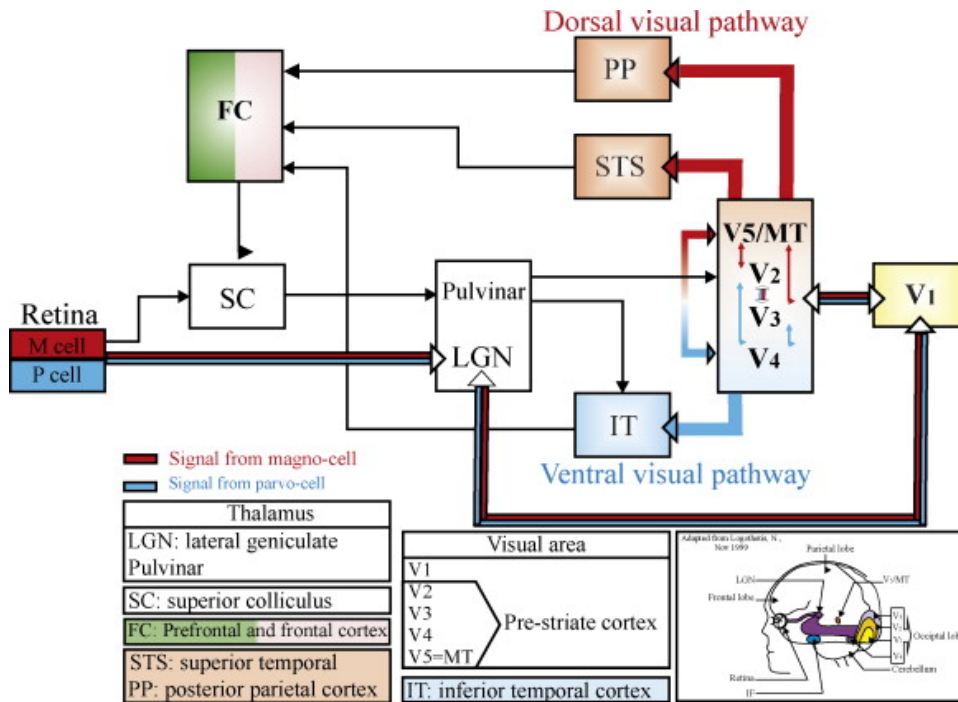


Figure 2.1.14: Simplified illustration of the visual pathway. After stimulation of the retina, *magnocellular* and *parvocellular* signals are conveyed through the LGN to the visual cortex (starting from V₁), and next they propagate to other brain areas. Figure taken from [188].

Fairly recently, a third K-pathway (after *koniocellular* neurons) was discovered which has properties that are roughly in between those of the M- and P-pathways in terms of speed and contrast perception, originating mainly from the blue cones, the K-pathway also carries blue and yellow color information.

Due to the cerebral cortex wrapping and to the respective areas sizes, V₁ area is the one better exposed for surface EEG recordings, while also V₄ is partially parallel and near to the skull.

At the moment most of the brain areas have been deeply studied with multiple techniques, from functional imaging to invasive electrodes recordings and stimulations, thus detailed maps are available. Anyhow the association between brain functions and their cortical areas has been debated for a long time and two main theories have been developed and sustained by evidences.

One theory, supporting high functional specialization, suggests the brain to have different modules that are domain specific in function, while the other theory, supporting distributive processing, proposes that the brain is more interactive and its regions are functionally interconnected rather than specialized [135].

Supporting the first theory there are evidences given by invasive single neuron recordings, showing highly selective neurons, firing only when a particular stimuli is received and only in particular brain areas. Also multi-sensory selective neurons have been found, firing in response to a “concept” than to a particular stimulus, named *grandmother cells*, since they may fire in the case of the vision of the face of the grandmother of the test subject, as well as reading her name or hearing her voice.

Nevertheless, strong evidence have been found supporting also the second theory concerning the plasticity of various cortical areas that are able to remap their functions, e.g. in the case of the somatosensory cortex, just some minutes after applying local anesthesia to a finger, the associated brain area start to remap over the nearby fingers [135].

Despite of this, regarding non-invasive Brain Computer Interfaces, where the spatial resolution of electrodes is quite coarse ² with respect to neurons populations, the available functional maps are still useful and used; e.g. the primary visual cortex is always found in the occipital lobe and various EEG signals associated to visual stimuli are better detectable by electrodes over that area.

Consequently for BCIs based on surface EEG recordings, a strict functional specialization of the cerebral cortex could be assumed, although it is probably a coarse approximation of the reality.

²Scalp EEG measures a space-averaged activity of about 10^8 neurons.

2.1.4 BRAIN RHYTHMS

As mentioned in Sec. 2.1.2 the synchronous spiking of a neuron population generate LFPs, as shown in Fig. 2.1.8, which are characterized by oscillations of the potential at a given frequency.

In a scalp EEG recording, various LFPs are summed together, giving a signal that is the summation of different locally synchronous neuronal activities; indeed some of the first observations performed on the EEG signals were related to their contents in the frequency domain.

The Power Spectral Density (PSD) of a generic EEG recording, in order to highlight the power distribution of its frequencies content, is shown in Fig. 2.1.15.

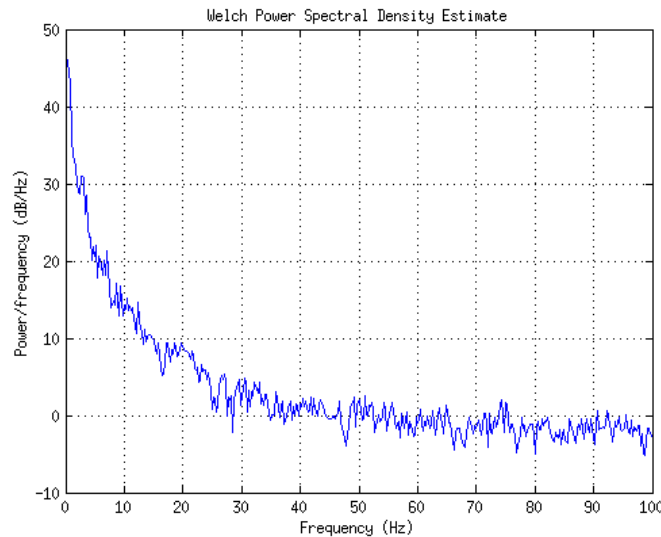


Figure 2.1.15: Example of the Power Spectral Density (PSD) estimate for 30 s of EEG data acquired at 256 Hz by one surface electrode over the visual cortex of an awake subject with eyes open at rest. PSD obtained with the Welch method using a 1024 points Hamming window [126]. The purpose of this image is just to show the power distribution across the frequency contents of an EEG recording.

The EEG is typically described in terms of its rhythmic activity and transient signals. Historically the rhythmic activity has been divided in different frequency bands which have been associated to different brain states by early researches in

the field. This band division is still commonly used and most of the early observations still holds, although more sophisticated analysis are commonly used today and various exceptions have been highlighted.

The main frequency bands associated to the human brain functions are:

- Delta (0 Hz to 4 Hz)
Rhythms in this band are mainly found frontally in adults and posteriorly in children and are characterized by high-amplitude waves. They are enhanced during non-REM sleep states.
- Theta (4 Hz to 8 Hz)
Rhythms in this band are mainly found in locations not related to task at hand and related to drowsiness or arousal.
- Alpha (8 Hz to 13 Hz)
Rhythms in this band are mainly found in the occipital region and are associated to relaxed/reflecting states and exhibit a strong amplitude increase when closing the eyes. In this band it is worth to mention also a particular rhythm called the “ μ rhythm” that is most prominent in the sensorimotor cortex when motor neurons are resting. Detecting its decrease in amplitude is therefore associated to the activation of motor neurons and this is commonly exploited in the implementation of Motor Imagery based BCIs as detailed in Sec. 2.3.2.
- Beta (13 Hz to 30 Hz)
Rhythms in this band are mainly found frontally, they are low-amplitude waves and are associated to active, busy, or anxious thinking, active concentration.
- Gamma (30 Hz to 100 Hz)
Rhythms in this band are mainly found in the somatosensory cortex and displays during cross-modal sensory processing or during short-term memory matching of recognized objects, sounds, or tactile sensations. This band is indeed of critical importance concerning the mechanism accounting for perceptual binding [174]; by synchronizing assemblies of neurons which

process various features of an object, gamma oscillations might allow ordinarily de-synchronous neurons to synchronize and therefore multiple their output on subsequent neurons. The gamma-binding hypothesis has been extended to include also binding across sensory modalities as audio-visual integration and even binding in a top-down sense, allowing for a more general object representation from memory rather than just bottom-up grouping [177].

2.2 EEG SIGNALS ACQUISITION

As previously mentioned, EEG recordings consist in measurements of voltage potential differences taken over the scalp of the subjects.

Although from such a description it may seem straightforward as using a multimeter and placing its electrodes on the scalp of a person, it is not. The EEG acquisition poses various challenges, some of whose are common to all biosignal recording systems, while some others are even more harder to tackle.

EEG measurements involve voltages at very low levels, typically ranging between $1\text{ }\mu\text{V}$ and $50\text{ }\mu\text{V}$, with high source impedances and superimposed high level interference signals and noise. Therefore the signals need to be amplified with relatively high gains to make them compatible with devices such as analog-to-digital converters (ADC) for computerized equipment.

The amplifiers used to measure EEG signals have to satisfy very specific requirements in order to provide a selective amplification of the physiological signal of interest rejecting superimposed noise and interferences. Amplifiers featuring these specifications, in general, are commonly referred as *biopotential amplifiers* [131].

Using commercial EEG acquisition devices part of the work has been tackled by electronic designers, but, despite of this, to the experimenter is still left the duty to choose between different settings and to comply with some guidelines for optimal performances.

In professional EEG devices, where the position of the electrodes can be decided by the experimenter, he/she is usually left with the duty to decide the *referencing system* to use, to mount the electrodes in order to obtain a good electrical connection with the skin, but also to choose the kind of electrodes. All of these

decisions, as well as instructions to the subjects to avoid some kind of behaviors (e.g. teeth grinding or muscle movements in general), may improve the recording quality, lowering the presence of artifacts.

2.2.1 SPATIAL CONFIGURATION OF ELECTRODES

Since the cerebral cortex is functionally divided in areas that are prevalently assigned to specific functions, the position of surface electrodes to record EEG signals is relevant accordingly to the underlying part of cerebral cortex.

To standardize the electrodes' positions over different experiments there exist some positioning standards, the most famous is the 10-20 system depicted in Fig. 2.2.1, adopted in 1958 by the International Federation in Electroencephalography and Clinical Neurophysiology. In this standard, distances between electrodes are given in percentages to adapt to subject differences in the skull dimension. It is called 10-20 system since distances between electrodes are either 10% or 20% of the total distances between the nasion (front) and inion (back) or right and left (right/left preauricular point).

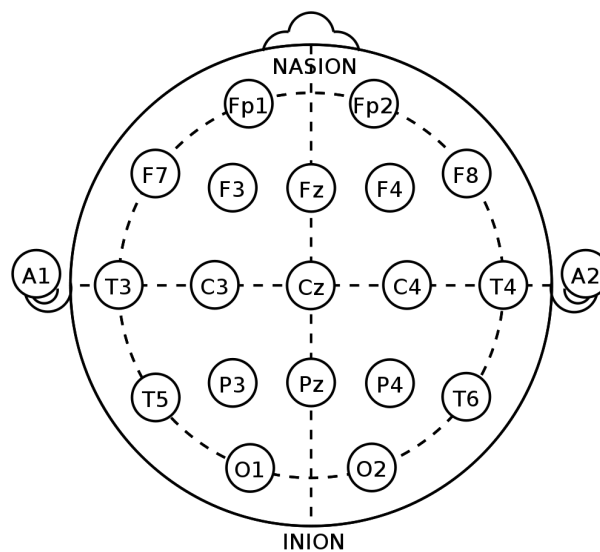


Figure 2.2.1: The 21 electrodes positions in the 10-20 standard system.

Every position is labeled by an initial letter indicating the main cortex area covered by the electrode (F: frontal, T: temporal, C: central, P: parietal, O: occipital),

the midline is labeled by a following “z” and the final number indicate the hemispheric deflection, where odd number are used for the left hemisphere, while even are used for the right hemisphere.

Other standards exists, mainly to extend the 21 electrodes 10-20 system to an higher number of positions; this allow standard recordings with more than 21 electrodes and/or a more precise positioning reference. High density EEG recordings may use even 256 electrodes or more, mounted on different scalp locations.

2.2.2 ARTIFACTS

In EEG signals recording, although the will is to record only voltage potentials generated by the neural activity, the obtained signals often (or always) contain the so called “artifacts” which are electrical signals detected along the scalp by the acquisition device, which originate from non-cerebral sources, superimposing over the signals of interest, and thus considerable as noises.

Such artifacts can be internal (generated in/by the subject body) as muscular activity, breathing, heart pulses, sweating, or can be also external, such as electromagnetic noise present in the recording environment. Some of the common EEG artifacts are shown in Fig. 2.2.2, highlighting also their high amplitude in comparison to the clean EEG.

Most of the internal artifacts not originating in the head, as well as external artifacts, are mostly discarded thanks to the *biopotential amplifiers* design in conjunction to a good referencing system, as will be detailed in Sec. 2.2.3 and a proper skin-electrodes impedance as will be detailed in Sec. 2.2.4.

Artifacts that will always be present in an EEG recordings, no matter which device and configuration is used, are the ones originating in the head of the subject, as muscular activity (e.g. eye blinks, eye movements, tongue movements, teeth grinding, etc.) or due to skin-electrode impedance changes due to sweating or electrode-skin movements. To reduce these kind of artifacts the only solution is to avoid the events originating them or otherwise to use post-processing software to detect and attempt to remove them.

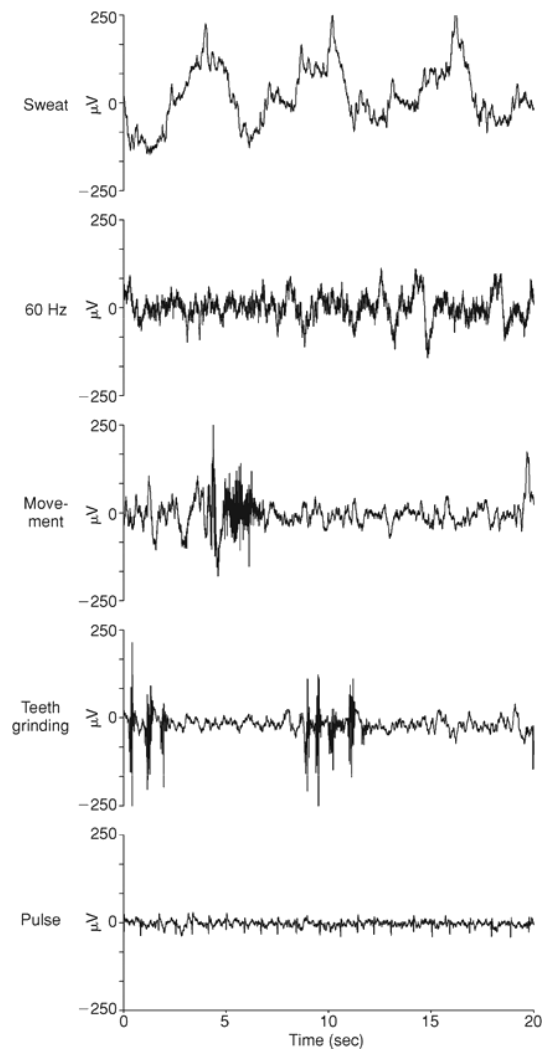


Figure 2.2.2: Common EEG Artifacts. From the top to the bottom, artifacts due to the: skin impedance change due to sweating; power line interference (American 60 Hz standard); a muscle movement; teeth grinding; heart beat pulse.

2.2.3 CHOOSING A REFERENCE

Since a difference between voltage potentials has to be measured, a reference point has to be identified. In other research fields it is a common practice to use as a reference the ground voltage, but in this case where very weak currents are generated on the scalp level and the voltage potentials are in the order of microvolts, any surplus of electrical charges (as static electricity) on the subject, with respect to the

ground, would hide completely the weak EEG signals and could saturate the amplifiers.

Moreover, measuring with respect to a “real ground”, would lead in the measurement all the fluctuations in voltage potentials generated from the whole subject body. A part from this, it is also widely known as a bad practice in biomedical engineering to directly connect a human body to the ground due to the risk of electricity discharge through the body of possibly bad-behaving electronic devices.

Furthermore, since high gain amplifiers have to be used in order to amplify the voltage potentials from tens of microvolts to several volts, electrodes and their wires can easily act as antennas, receiving whatever is present in the electromagnetic spectrum, adding it as noise over the signals of interest.

Consequently, a preferable solution would be to measure differences between two different electrodes positioned on the body, using this method all the artifacts present at both the electrodes’ sites would elide and this would hold also for static electricity.

Indeed, a typical configuration for the measurement of biopotentials consists of at least three electrodes E , R and G . Two of them E and R are used to record the signal as the potential difference between them with respect to the common ground G . The desired biopotential signal is named *differential signal*, while the signal appearing between the inputs and ground is named *common mode signal*. The Common Mode Rejection Ratio (CMRR) of an amplifier is defined as the ratio of the differential mode gain over the common mode gain and thus it identifies the amplifier ability to remove from the output amplified signal the interference signal common to the input electrodes [131].

In practice, for a single electrode EEG recording, an electrode E is placed on a scalp position where the EEG has to be acquired and its voltage potential is measured with respect to a ground electrode G that could be placed anywhere on the body (which is commonly placed on the scalp too), concurrently another reference electrode R is placed where neural activity should not be detectable, but where possibly other internal artifact signals could be read (as the earlobes or the nose). Having two signals, one given by the difference between $E - G$ and one by the difference between $R - G$, a further difference is taken between them, removing the ground effect, but also some artifacts common to all of them.

The resulting signal therefore reflects the ongoing neural activity under the E electrode position with respect to R .

High density EEG recordings may use up to 256 electrodes or more, but for Brain Computer-Interfaces simpler and easier to use set-ups are commonly used, from few electrodes up to few tens.

For multiple electrodes recordings, only one electrode is commonly used as ground, but two different referencing system could be used. Referencing multiple electrodes to the same reference electrode as shown in Fig. 2.2.3 is referred as *unipolar derivation* recording, while *bipolar derivation* recording could also be performed using different couples of electrodes as shown in Fig. 2.2.4.

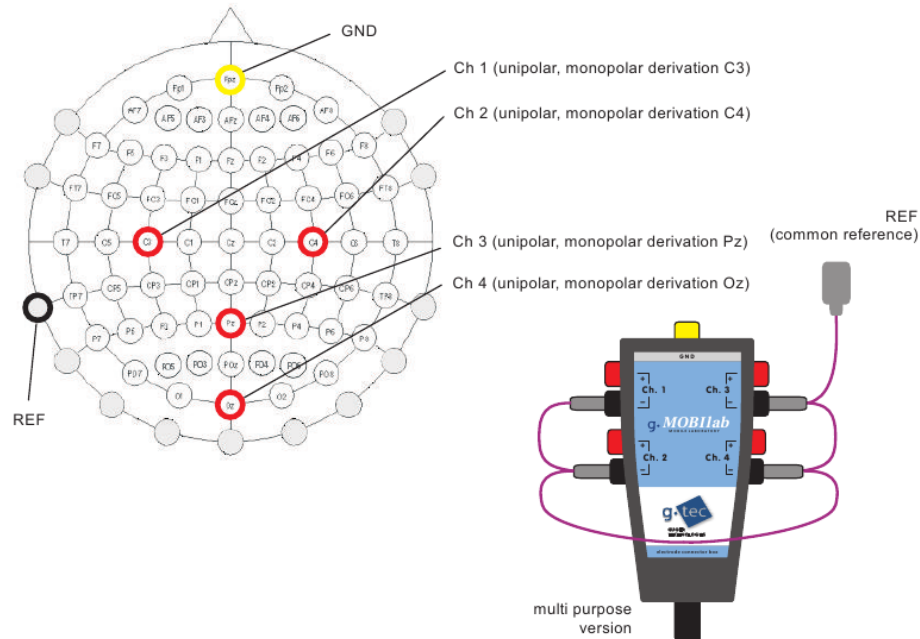


Figure 2.2.3: Example of a possible configuration of the g.tec g.MOBilab+ multipurpose version to perform an *unipolar* EEG recording. Detailed description of the device will be given in Sec. 5.1.2. Figure adapted from [65].

The main issue of the latter approach is given by the fact that the read signal is influenced by the neural activity occurring under both the electrodes and there could be the unlucky case in which if an electrode read a positive potential and the

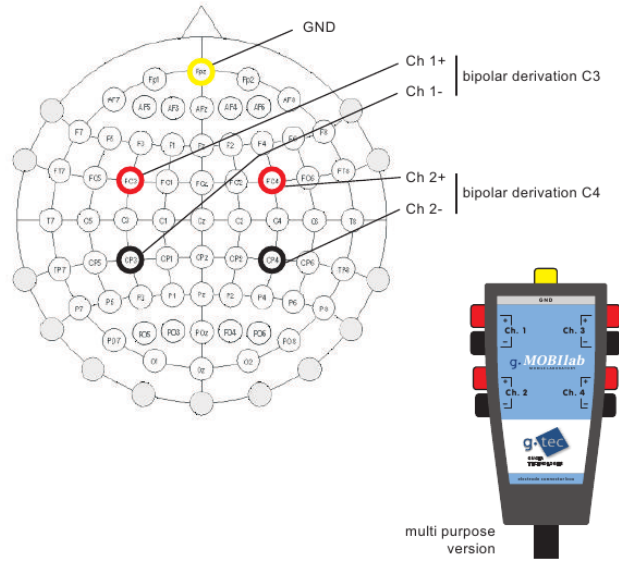


Figure 2.2.4: Example of a possible configuration of the g.tec g.MOBIlab+ multipurpose version to perform a *bipolar* EEG recording. Detailed description of the device will be given in Sec. 5.1.2. Figure adapted from [65].

other a negative one, they may elide each other, giving a flat signal despite of strong (and opposite sign) activities may be occurring on both sites.

In fact, one of the most used referencing system is the *unipolar derivation*, also due to the fact that with an unipolar recording it is just a matter of subtracting different channel signals to apply a virtual reference and obtain the corresponding bipolar one.

2.2.4 ELECTRODE IMPEDANCE

Although the human skin can conduce electricity, it is not a perfect conductor and in particular, it is not straightforward to achieve a good connection between the skin and the electrodes.

Actually electrodes connection quality to the skin scalp, can be assessed and measured with hardware tools such impedance-meters. Some acquisition devices have embedded meters, while others need the operator to check electrodes' impedance before connecting them to the device.

Impedance extends the concept of resistance to Alternate Current (AC) cir-

uits, and possesses both magnitude and phase, unlike resistance, which has only magnitude. When a circuit is driven with Direct Current (DC), there is no distinction between impedance and resistance, but EEG signals can not be thought as DC signals since they are varying in time with frequencies of interest spanning from about 0.1 Hz to 50 Hz or for some applications from 0.1 Hz to 100 Hz.

In a lot of works using EEG signals acquisition, it is mentioned that the impedance of the electrodes is checked and kept lower than 5 k Ω ; being the impedance Z a complex number, representing in simple words the resistance with respect to different frequencies, this statement is quite misleading from a theoretical point of view.

From a practical point of view, since the EEG signals are limited to a relatively narrow-band from 0.1 Hz to 100 Hz (and are more commonly taken into account between 0.1 Hz to 50 Hz), the impedance is commonly checked at a single frequency and thus only the “resistance at a given frequency” is reported.

Also for the device used in this work, all the impedance measurements are taken at a single frequency and in particular at 10 Hz, accordingly to the impedance-meter specifications, and will be reported in kilo-ohm. Since the impedance meter and the acquisition device are designed to be used together and from the device’s manual [64] the electrodes impedance is requested to be kept under 5 k Ω , the requested procedure will always be followed.

Skin-electrode impedance is very important for biosignal acquisition in general, but in particular for EEG recordings [131], since the rejection of the common mode signal is a function both of the amplifier CMRR and of the source impedances z_{ea} and z_{eb} (with reference to Fig. 2.2.5). In Eq. 2.1 is shown the relation for a simple biosignal amplifier between its gain G , its input impedance Z_I , the differential signal V_s , the common mode signal V_c and the electrodes impedances z_{ea} and z_{eb} .

$$V_{out} = GV_s + \frac{GV_c}{CMRR} + GV_c \left(1 - \frac{Z_I}{Z_I + Z_{ea} - Z_{eb}} \right) \quad (2.1)$$

In ideal conditions the electrodes impedances should be equal and the CMRR should be infinite, but unfortunately this is not the case for real devices. With a

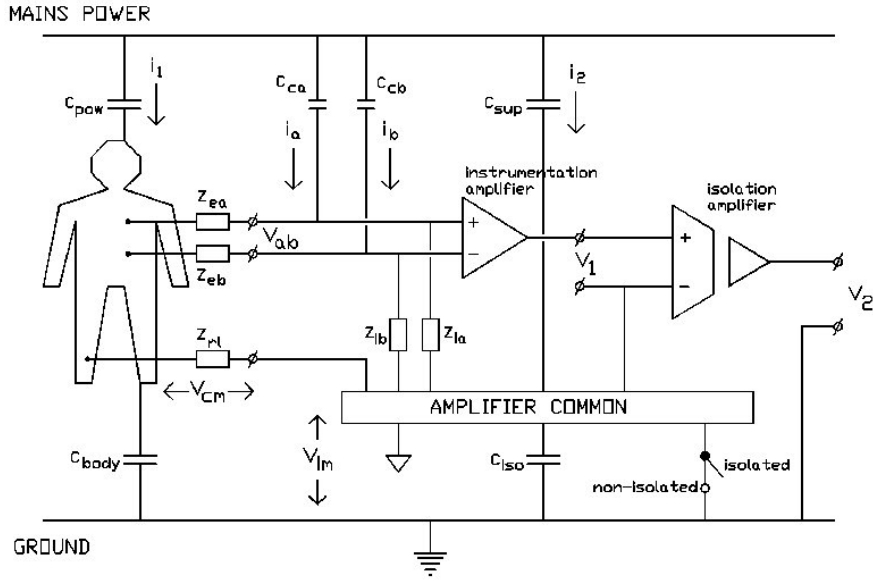


Figure 2.2.5: Schematic representation of a generic simple bioamplifier showing the main connections and source of interferences. All the depicted capacitors represent the actions of ambient interferences as the power line noise and do not represent actual connections. The figure is assuming an electrocardiographic application, but the same scheme is valid also for EEG amplifiers a part from the electrodes positions on the body. Figure taken from [186].

finite CMRR the common mode signal would not be completely rejected and the same would happen with unbalanced source impedances.

Since source impedance unbalances of $5\text{ k}\Omega$ to $10\text{ k}\Omega$, mainly caused by skin-electrode connections, are not uncommon and sufficient rejection of frequency interference requires a minimum CMRR of 100 dB, the input impedance of the amplifiers are commonly in the order of $10^9\text{ }\Omega$ at 50 Hz (or 60 Hz) and biopotential amplifiers provide commonly a CMRR in the order of 100 dB to 200 dB [131].

Therefore lowering the source impedance under $5\text{ k}\Omega$, means to avoid unbalances greater than that and therefore granting a more efficient rejection of the common mode signal, that in the case of EEG recordings, a part from the external artifacts, comprises most of the internal artifacts not originating from the subject head.

To reduce the electrode impedance with the skin, specific conductive gels are used and moreover the skin under the electrode should be prepared removing the

natural grease that is commonly over it with some alcohol. Most of the available gels are indeed abrasive in order to let the operator further prepare the skin rubbing the epiderm in order to remove the layers of dead cells that may increase the impedance with the electrode.

Actually not all the electrodes and amplifiers have the same requirements and specifications, but the aforementioned details holds in general for most of the state-of-the-art bioamplifiers and electrodes. Despite of this, for some applications, as for BCI applications, researchers are working towards new devices in order to avoid the use of skin preparation and abrasive gel in favor of the user comfort and to renounce to some features and flexibility in favor of cheaper and easier to use devices.

2.2.5 KINDS OF ELECTRODES

Very different kind of electrodes exists and can be differentiated mainly in four main categories: classical gel based contact electrodes, water based contact electrodes, dry contact electrodes and capacitive non-contact electrodes.

The materials used to build electrodes can have different impacts on the recording signals, in particular the electrodes are not picking up electron flows, but ions concentrations [100], thus the chemical reactions happening at the skin-electrode contact can strongly influence the acquired signal for contact electrodes.

Commonly used electrodes for EEG recordings can be made of different conductive metals with different properties [176]: Au, Sn, Ag, sintered Ag/AgCl, Pt, stainless steel, etc. Anyhow the most used on professional devices are Ag/AgCl electrodes and Au plated Ag electrodes. Both of them need the use of a conductive gel or paste and are contact electrodes.

An EEG electrode is a transducer that senses ion distribution on the surface of tissue, and converts the ion current to electron current. An electrolyte gel or paste is commonly placed on the side of the electrode that comes into contact with the skin, while the actual electrode consists of conductive metal attached to a lead wire connected to the amplifier.

A chemical reaction occurs at the interface between the electrolyte and the electrode. Current crosses the skin-electrode interface as the atoms in the electrode oxidize to form cations and electrons. The cations are discharged into the elec-

trolyte, and the electrons carry charges through the lead wires. Similarly, the anions in the electrolyte travel toward the interface to deliver free electrons to the electrode. A voltage known as the *half-cell potential* develops across the interface due to an uneven distribution of anions and cations and it appears as a DC offset in the recording which is strongly dependent on the material of the used electrode.

As previously mentioned a very popular electrode is the silver/silver chloride Ag/AgCl one because of its very low *half-cell potential* of approximately 220 mV and its ease of manufacturability. Anyhow they must be chlorinated and due to their easy performance degradation in time they are often used as disposable electrodes or otherwise Au plated electrodes are used instead. Au plated electrodes provide an higher impedance [100] and are not suited for low frequency recordings under 0.1 Hz (often referred as DC recordings), but they are maintenance free and long-lasting. Indeed in this research work Au plated electrodes will be used for the g.MOBILab+ acquisition device described in Sec. 5.1.2.

Au and Ag/AgCl electrodes are both non-polarized electrodes and they allow current to pass across the interface between the electrolyte and the electrode. Non-polarized electrodes are better than polarized electrodes in terms of their rejection of motion artifacts since they may charge up the capacitance from the electrolyte and electrode interface. In this regard, Sn or stainless steel electrodes may cause lower signal quality due to polarization noise.

Current can pass from an electrolyte to a non-polarized electrode, while polarized electrodes act more like a capacitor and current is displaced but does not move freely across the electrolytic interface.

Contact-electrodes which do not need gels or skin preparation also exist and are known as dry electrodes [204], although they are known to be more noisy [116, 117] than gel based ones, they are being investigated in particular for the BCI field where easy of use for end user applications is essential. Commercial dry electrodes EEG acquisition devices are already available, both as commodity devices^{3 4} and professional devices⁵ aimed mainly to BCI applications.

Non-contact electrodes also exist and are indeed based on the idea of being one

³<http://www.neurosky.com/>

⁴<http://www.koenenco.nl/en/product/eeg-headset/>

⁵<http://www.gtec.at/Products/Electrodes-and-Sensors/g.SAHARA-Specs-Features>

plate of a capacitor, while the surface of the skin act as the other plate. Those kind of electrodes are not suited at all for low frequency recordings (where for low in this case it is meant several Hertz), since acting as capacitors they completely cut lower frequencies. Despite of this they are under investigation for the production of commodity devices not requiring any skin preparation, nor skin contact, which could be extremely interesting for BCI applications [139].

Another main difference between electrodes is if they embed a pre-amplifier or not; electrodes embedding in their body a pre-amplifier to send along their cable a stronger signal, less prone to electromagnetic interferences, are known as active electrodes, while passive electrodes are the one not embedding any electronics.

2.3 BCIs MODALITIES

In literature, different BCI modalities have been successfully adopted and differentiate between them accordingly to the kind of underlying brain process which is used to detect features associated to brain states. The most popular are based on the Event Related Potentials (ERP), on the Event Related Synchronization / Desynchronization (ERS/ERDS) or Motor Imagery (MI) and on the Steady-State Visually Evoked Potentials (SSVEP).

2.3.1 EVENT RELATED POTENTIALS AND THE P300

In the context of non-invasive EEG recordings, Event Related Potentials, or ERPs, are transient voltage potential variations measurable on the scalp surface, interpretable as a brain response to an event and thus event-related.

ERP waveforms have an amplitude such that it is quite hard to identify them inside the recorded EEG signal, since it is containing the event-related waveforms superimposed to many other signals, given by event-unrelated neuronal activity.

In Fig. 2.3.1 is represented a typical ERP experiment, where multiple stimuli are presented multiple times to the user in order to elicit multiple ERP waveforms which, as can be seen by visual inspection, are not easily detectable inside the raw EEG recording.

In order to extract from the raw EEG the event related waveform, it is performed

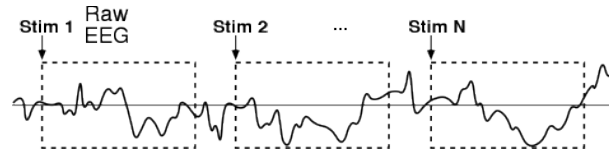


Figure 2.3.1: Multiple presentation of stimuli in a generic Event Related Potentials experiment. Figure courtesy of <http://erpinfo.org/>.

a procedure of averaging in the time domain, time-locking the different recorded epochs⁶ on the moment the stimulus was presented. In this manner, every event-unrelated component of the raw EEG should elide in the average operation, since it is not present in every epoch after the same amount of time from the stimulus presentation event, while on the contrary, all the event-related components will sum as depicted in the right of Fig. 2.3.2.

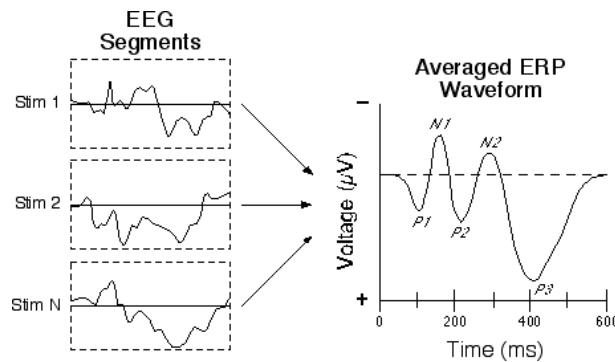


Figure 2.3.2: Time-locking and averaging of the different epochs of a raw EEG recording in a generic Event Related Potentials Experiment. Figure courtesy of <http://erpinfo.org/>.

There exist various known ERP waveforms associated to different kind of events, stimuli and conditions, but the most famous in the context of ERP based BCIs is the P300 or P3 waveform which is a positive waveform peaked between 300 ms to 400 ms after the stimulus presentation [51].

The P300 is considered to be an endogenous potential, as its occurrence links not in particular to the physical attributes of the presented stimulus, but mainly to the user's reaction to it. In particular, the P300 is thought to reflect processes

⁶An EEG epoch is defined as a temporal slice of a longer EEG signal.



Figure 2.3.3: The letters matrix displayed by the OpenVibe software P300 speller scenario. Figure courtesy of <http://openvibe.inria.fr/>.

involved in stimulus evaluation or categorization. Indeed it is usually elicited using the *oddball* paradigm, in which low-probability target items are mixed with higher-probability non-target items. In simple words, the P300 waveform is more likely to occur after a rare and relevant stimulus is presented to the subject; to be relevant a stimulus has to be “searched” by the subject. As an example, the subject may be instructed to count how many times a particular picture is shown in a sequence of different pictures visualizations. Every time the “searched” picture is shown a P300 waveform should appear post-stimulus in the subject EEG recording.

A classical BCI application using the P300 ERP component is the so called P300 matrix speller, which was first described by Farwell and Donchin in 1988 [51]. It implements the *oddball* paradigm in order to let the BCI user to communicate to a computer the will to select a letter or a symbol from a list, simply gazing at it. In other words the P300 speller is a BCI allowing its users to write without using their limbs or any other muscular activity a part from their eyes⁷.

More precisely, the usage of a typical P300 Speller BCI can be divided in three main phases:

⁷There is a debate on the need of eye’s muscles functionality to utilize a P300 based BCI, but have been recently demonstrated that to gaze at the stimulus (and not to only covertly attend to it) is quite essential [18].

ACQUISITION OF TRAINING DATA

In this phase the user is requested to attend to a given letter or symbol, displayed in a matrix as the one presented in Fig. 2.3.3. Then a column or a row of the matrix is lit randomly and repetitively in order to lit every column and every raw multiple times (since as shown in Fig. 2.3.2 multiple epochs are needed for the time-lock average). All the EEG epochs recorded after the lit of every raw or column are averaged with the ones related to the same raw and column.

When the column or the raw containing the target letter is lit, a P₃₀₀ waveform should be elicited and it should be recorded in the respective epochs and “amplified” by averaging them together. This can be repeated multiple times in order to collect more data.

Since the letter the user was starring at, in this phase, is known, at the end of this procedure, different EEG averaged epochs will be obtained and moreover will be known which of them should contain a P₃₀₀ waveform corresponding to the attended letter.

CLASSIFIER TRAINING

A classifier is then trained in order to learn to differentiate between averaged epochs containing a P₃₀₀ waveform, corresponding to the attended letter and epochs not containing it.

In this phase to have enough data to perform the training is extremely important to obtain high classification performances.

FREE USE

In the last phase, technically similar to the first one, the user is eventually free to choose which letter or symbol to attend and therefore untagged epochs are produced every time a row or a column is lit.

The epochs regarding the same row or column are averaged and the classifier has to identify which of the averaged epochs contains a P₃₀₀ waveform corresponding to the attended letter.

When both a row and a column are identified, a response on the most probable letter or symbol the user was attending can be given.

One of the major disadvantage of the ERP based BCIs is the need of the time-lock averaging operation, since this force to have a precise stimulus presentation time synchronization with the EEG recording, but also because it requires multiple stimuli presentations to collect enough EEG epochs.

Analyzing brain states that correspond to ERPs on a single trial basis is a challenging problem due to the high trial-to-trial variability and the unfavorable ratio between signal (ERP) and noise (artifacts and neural background activity), nevertheless, recently the use of advanced mathematical tools, reported interesting results also in single-trial ERPs identification (although with some limitations) [13].

2.3.2 EVENT RELATED SYNCHRONIZATION/DE-SYNCHRONIZATION

As the ERP based BCIs exploits event-related transients in the temporal domain of the EEG signals, BCIs based on the Event Related Synchronization (ERS) and Event Related De-synchronization (ERD) modality exploits event-related increase and decrease of power in the frequency domain.

The terms ERS and ERD are due to the fact that an increase or a decrease in power at specific frequency or band in an EEG signal, corresponds to a synchronization or de-synchronization of a corresponding population of neurons.

A popular BCI based on this modality is able to detect actual or, more interestingly, imagined, movements of the limbs. Commonly left and right hand movements have been a popular choice, but also tongue and feet have been used [182]. This kind of BCIs, using the ERS/ERD modality are often called Motor Imagery (MI) BCIs [144] since the subject is commonly asked to imagine a movement in order to instruct a command.

MI BCIs works thanks to the fact that, as mentioned in Sec. 2.1.4, large populations of neurons in the motor cortex area synchronize at a given frequency, in the 8 Hz to 13 Hz range, when the subject is at rest. Patterns in this band, over the motor cortex, referred as μ -waves or μ -rhythms, are suppress (due to a de-synchronization) by the subject when he or she performs a motor action. Mental imagination of movements seems to involve similar brain regions and functions which are involved in programming and preparing of movements. In fact, it is ac-

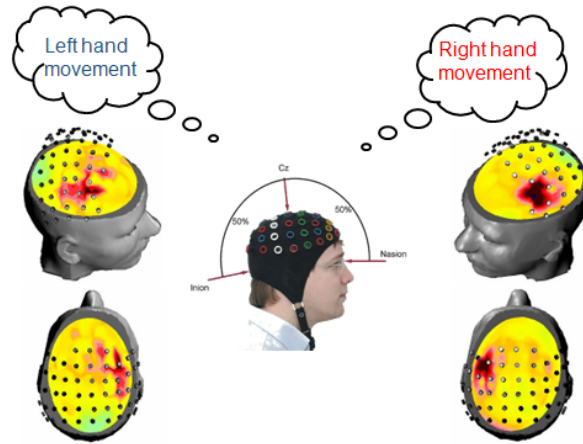


Figure 2.3.4: Graphical representation of a MI based BCI, where different cortical activation patterns can be observed, plotted on the corresponding cortical areas, according to the kind of imagined movement. Figure courtesy of <http://gtec.at/>.

cepted that the main difference between performing a movement and imagining it, is that in the latter case the execution is blocked at some cortico-spinal level [144], but the motor cortex activity share common neural mechanisms for imagination, preparation and actuation of movements [41].

Being each side (left/right hemisphere) of the motor cortex controlling the contralateral side of the body, a movement of the right hand will lead to a de-synchronization of the μ -rhythm in the left hemisphere motor cortex, in particular over the area controlling the hand. As mentioned before, simply imagining the same movement, would lead to a similar de-synchronization as depicted in Fig. 2.3.4. Moreover has been demonstrated that the subjects can learn to “better imagine the movement” in order to increase the de-synchronization and thus to augment the detection efficiency and reliability of the BCI.

To increase the detection accuracy, algorithms to increase the variance between the different conditions (e.g left/right imagination) are commonly used, as the popular Common Spatial Pattern (CSP) [152], which will be described in detail in Sec. 4.3.1. To discriminate between the different conditions machine learning algorithms are commonly used as the LDA (Linear Discriminant Analysis) or SVM (Support Vector Machines).

As for the P300 speller, the usage of a typical Motor Imagery BCI can be divided in three main phases:

ACQUISITION OF TRAINING DATA

In this phase the subject is instructed to imagine for example the movement of the left hand and the right hand in different trials while EEG signals are recorded from electrodes positioned over the motor cortex.

In this phase various epochs are recorded for the different conditions to be used in the next phase (e.g. left hand movement and right hand movement).

CLASSIFIER TRAINING

In this phase the previously recorded epochs are filtered in the band of interest (commonly 8 Hz to 13 Hz) and algorithms as CSP may be applied in order to compute linear filters able to increase the variance between the different conditions.

After this pre-processing operation a classifier can be trained to discriminate between the two conditions, using as features the power of the signal in the selected frequency band.

FREE USE

When the classifier is trained, the free use of the BCI is possible and the user imagining hand movements can instruct commands for example to control a wheel chair or an avatar in a virtual environment [109].

As for most of the ERP based BCIs, also MI based BCIs commonly need the use of machine learning and thus a “calibration” procedure has to be performed for each subject before the actual use of the BCI. Furthermore also the same subject is commonly requested to perform the calibration for each session, since electrodes position may slightly vary requiring different spatial filters.

Interestingly, in the case of MI based BCIs, in contrast to the P300 based BCIs case, no stimuli has to be presented and no time-locking operations in the time do-

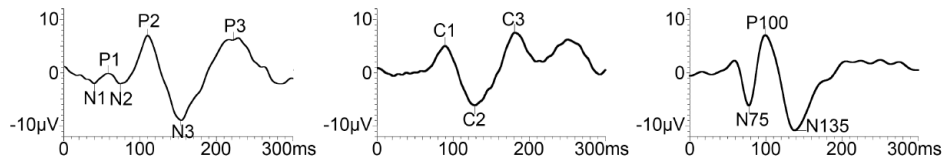


Figure 2.3.5: From left to right, a VEP elicited by a flash of light, one elicited by onset/offset of a pattern and one elicited by a pattern reversal. Figure taken from [10], originally adapted from [138].

main has to be performed, consequently a synchronization with the EEG recording is not needed. Probably this characteristic lead MI one of the preferred approaches followed for the implementation of practical BCIs.

2.3.3 STEADY-STATE VISUALLY EVOKED POTENTIALS

Any change in the visual field content can elicit in an human brain what is called a transient visual evoked potential (tVEP or simply VEP).

A VEP is characterized by a waveform similar to the ERPs described in Sec. 2.3.1 and again a common technique to “extract” it from the unrelated brain activity is to average over several trials in the time domain, time-locking the EEG epochs at the stimulus presentation time.

To elicit a VEP, a common method is to show to the subject some stimulus to generate a sudden change to the content of the subject’s visual field. Common used stimuli are a flash of light, or a colored shape appearing and disappearing on a screen, or otherwise a pattern like a checkerboard reversing its black and white boxes. According to the chosen stimulation, very different waveforms could be obtained, as shown in Fig. 2.3.5.

VEPs have been used in the past and still today, as a means of studying the Human Visual System (HVS) functioning, but also in order to diagnose cognitive and vision disorders [138].

In the sixties, Regan [154], started to investigate the use of long visual stimuli trains of sinusoidally modulated monochromatic light in order to investigate the EEG waveforms given by a series of VEPs. The response given by the presentation of these kind of repetitive visual stimulation, averaged in the time domain over

multiple trials, looked like a sinusoidal waveform oscillating at the same frequency of the stimulus. These EEG waveforms, having a lower amplitude than transient VEP, but being constant in their frequency and phase, were named steady-state VEP or SSVEP.

Being the SSVEP waveforms characterized by a quasi-sinusoidal waveform of a fixed frequency buried in the event-unrelated brain activity, from a signal analysis point of view, the presence of a SSVEP response could be assessed much more easily in the frequency domain after a Fourier Transform of the EEG signal.

Most of the SSVEP based BCIs are implemented showing to the user a set of objects flashing (or reversing their pattern) at a specific frequency. Every object named “target”, is commonly flickering at a different frequency (*frequency tagging* method) and a command can be associated to each of them. When the user focus his/her attention (overtly or covertly [192]) on the desired target, the measured brain’s activity frequency components increase for the target’s frequency and its harmonics.

This particular BCI modality, being the one chosen in this work will be addressed in depth in Chap. 4. Anyhow, as for other modalities, the usage of a generic SSVEP based BCI is here reported in its three main phases:

ACQUISITION OF TRAINING DATA

In this phase the subject is instructed to attend a particular target between the presented ones, or to not attend to any of them. In the meanwhile the EEG is recorded and thus various epochs are saved, where every epoch is known to correspond to a particular target, flickering (or changing) at a specific frequency.

CLASSIFIER TRAINING

During the training phase various kind of features could be extracted from the recorded epochs for classification. One of the simplest approaches is to compute the power of the recorded signal in a narrow-band (between 0.2 Hz to 1 Hz) centered around each of the stimulation frequencies and training a classifier to recognize which target is being attended using a feature vector composed of the power values in the evaluated narrow-bands.

More sophisticated approaches may use CSP algorithms to find linear combinations between the signals coming from different electrodes, in order to increase the variance between conditions, as will be addressed in Sec. 4.3.1. Otherwise, signal models of the SSVEP response could be used to compute a sort of signal to noise ratio (SNR) between it and the background neural activity, as detailed in Sec. 4.3.1.

Using this kind of SNR, the use of a classifier could also be avoided and the simple identification of the higher SNR among the ones corresponding to the different target frequencies could be computed to identify the attended target. A simple threshold could be further adopted to avoid false positives when the user is not attending any target.

FREE USE

Once the classifier has been trained, the user can start to use the BCI gazing at the different flickering targets to issue the commands associated to them.

In this phase the same feature extraction algorithms have to be used as during the classifier training.

2.4 BCIs CATEGORIES

As introduced at the beginning of this chapter, the term BCI has been used to define a lot of different kind of systems, based on the presented modalities, but also based on different working principles.

As already mentioned, in this work, a broad definition of the term BCI will be adopted, but for clearness and to highlight the different paradigms implemented, will by adopted also a particular categorization presented in [201, 203] which in my opinion is spreading in the community.

According to [201], BCIs can be divided into three main categories with smooth boundaries: *Active*, *Reactive* and *Passive* BCIs. Every modality (as the ones presented in Sec. 2.3) may fall inside one of them, accordingly to how it is used, although not every modality is suitable for every category.

Other kind of categorization could also be found in different research works as the differentiation between *Synchronous* and *Asynchronous* (or *Self-paced*) BCIs or

between *Dependent* and *Independent* BCIs.

2.4.1 ACTIVE BCIs

Active BCIs, sometimes referred as endogenous BCIs, derive their outputs from brain activity which is directly consciously controlled by the user, independently from external events, for controlling an application.

The MI based BCIs are an example of a classical *Active* BCIs. The brain activity being detected is consciously controlled by the user and no external stimuli has to be presented to evoke it.

2.4.2 REACTIVE BCIs

Reactive BCIs, sometimes referred as exogenous BCIs, derive their outputs from brain activity arising in reaction to external stimulation, which is indirectly modulated by the user for controlling an application.

This is the case for P300 and SSVEP based BCIs, where the brain activity being detected is evoked by an external stimulus and only modulated by the user e.g. gazing a particular target to instruct a command.

2.4.3 PASSIVE BCIs

Passive BCIs derive their outputs from arbitrary brain activity without the purpose of voluntary control and their main application is for enriching a human-computer interaction with implicit information.

In the field of *Passive* BCI, the user is commonly not instructed/trained to modify its brain activity in order to let the interface to execute a command. The user can be unaware of the BCI itself and the information extracted from the brain activity is commonly used as a secondary communication channel in Human-Machine Systems (HMS), where the primary channel could be implemented with an “ordinary” interaction device.

Passive BCIs can therefore be used to extract from the EEG signals cognitive states or emotional states of an user engaged in a task [201] and in fact most of the BCIs able to detect affective states of the user, commonly named Affective BCIs (aBCIs), are *Passive* BCI.

2.4.4 DEPENDENT VS INDEPENDENT BCIs

In some research works BCIs could be also divided in *Dependent* BCIs and *Independent* BCIs. A *Dependent* BCI does not use the brain's normal output pathways of peripheral nerves and muscle as a signal source, but need some nerves and muscles to be functioning, e.g. to shift the eyes gaze. On the other hand an *Independent* BCI is able to be operated also by completely paralyzed subjects. This categorization has its roots in the research works aiming to provide BCIs as a means of communication for locked-in patients.

2.4.5 SYNCHRONOUS VS ASYNCHRONOUS BCIs

Another frequent categorization is between *Synchronous* BCIs and *Asynchronous* BCIs. In a *Synchronous* BCI, the system evaluate the user brain activity in system-defined time-windows in which the user is supposed to issue a command. On the other hand in an *Asynchronous* BCI, the system is continuously analyzing the user brain activity, while the user is free to issue a command anytime he/her wants.

Asynchronous BCIs are therefore harder to implement since the system has not only to discriminate between the available commands, but it has also to discriminate between the condition in which the user do not want to issue any command and when the user is trying to do it. *Asynchronous* BCIs are often referred also as *Self-Paced* BCIs.

2.4.6 COMPARING DIFFERENT BCIs

To compare different BCIs is commonly used the classification accuracy of the system between the available commands, which is defined as the probability P that the system correctly classify the user intent.

As a more comprehensive measure is often used also the Information Transfer Rate (ITR) value, in order to evaluate not only the classification accuracy, but how much information can be communicated in a period of time. Actually it can be used only for *Active* or *Reactive* BCIs, since it assumes the user being voluntarily instructing the system to execute a command.

The ITR value has been introduced in order to take into account both the speed of a BCI in detecting a user command and its accuracy in detecting the correct

command [114]. Its measuring unit is bit s^{-1} , although it is more commonly used as bit min^{-1} . ITR is a standard measure for communication systems based on the Shannon's information theory, which takes into account the accuracy, the number of possible selections and the time required to make each selection. The bitrate in most general form can be reduced to the mutual information between the actual and expected classification of the system. Nykopp's definition of the bitrate follows from [10]:

$$B = (X; Y) = H(Y) - H(Y|X) \quad (2.2)$$

$$H(Y) = - \sum_{j=1}^M p(y_j) \log_2 p(y_j) \quad (2.3)$$

$$p(y_j) = \sum_{i=1}^N p(x_i) p(y_j|x_i) \quad (2.4)$$

$$H(Y|X) = - \sum_{i=1}^N \sum_{j=1}^M p(x_i) p(y_j|x_i) \log_2 p(y_j|x_i). \quad (2.5)$$

where X represents the expected outcome, while Y the actual one; $p(x_i)$ is the a priori probability that the i^{th} symbol is expected, $p(y_j)$ is the probability that any signal is classified as the j^{th} one, while $p(y_j|x_i)$ is the probability that the system classifies a signal as the j^{th} symbol, given that it is actually the i^{th} . I is the mutual information, while H is the entropy.

Despite of this, in order to simplify the computation of the ITR, most of the research works published so far adopt some assumptions [10, 114]:

- It is assumed that all the symbols have the same a priori probability:

$$p(x_i) = 1/N$$

- That the classifier accuracy P is the same for all symbols, thus for $i = j$:

$$p(y_j|x_i) = P$$

- That the classification error $1 - P$ is equally distributed amongst all remaining symbols:

$$p(y_j|x_i) = \frac{1 - P}{N - 1}$$

Adopting these reasonable assumptions, as described also in [164], the number of bits B transmitted in a time window can be computed as:

$$B = \log_2 N + P \log_2 P + (1 - P) \log_2 \frac{1 - P}{N - 1} \quad (2.6)$$

where N is the number of possible symbols (or commands), and P is the probability that the symbol is correctly detected. The bit rate as bit min^{-1} , can then be computed by dividing B by the time window duration in minutes.

As an example, SSVEP based BCIs as a peak performance in optimal conditions can reach a transfer rate of 68 bit min^{-1} [207].

For *Passive* BCIs, being impossible to clearly determine if a “target” is correctly identified, the evaluation is commonly performed by other means e.g. using questionnaires or other forms of evaluations borrowed from the Human-Computer Interaction (HCI) research field. The impossibility to compare *Passive* BCIs to the other BCIs in terms of their ITR is also another reason why some researchers would prefer to not consider them in the BCI definition.

3

BCIs in Virtual Reality and Computer Games

The main goal of the research in the field of Brain-Computer Interfacing, as already mentioned, was initially to provide a new means for communication and control for mobility impaired subjects [197]. In the last decade various results have been obtained in this direction, providing BCI operated spellers, wheelchairs drivers and also prosthetic limbs controllers [113], which in conjunction to the production of commodity EEG devices will soon lead to a wider spread of commercial applications in the field.

More recently new applications have been envisioned, and in some cases implemented, to exploit BCI technologies also to provide new communication and control tools for healthy users, in particular in the areas of multimedia and entertainment [136, 146].

The use of BCI technologies in conjunction with Virtual Reality (VR) and/or Computer Gaming may have interesting potentials under two different point of

views [109]: from the VR community, BCIs are perceived as new input devices that may provide new tools to interact with Virtual Environments (VE) [98], while on the other side, from the BCI community point of view, VR can provide richer and more motivating feedbacks to the users than simple 2D representations, reducing the time needed to learn to use the BCI, as well as increasing the mental states classification performances ¹ [165]. Moreover from the BCI community point of view, VR could be used also as a safe and cost effective approach to test BCI aimed for the real world use [102].

Some researchers consider non-invasive BCIs still to slow and unreliable to provide new means of interactive controls to healthy users able to substitute ordinary devices [113, 146], as keyboards, mice, etc. Despite of this, non-invasive BCIs, leaving to the healthy users the ability of using their own limbs to operate also ordinary interaction devices, could be able to integrate them, rather than substitute them [108]. This approach has indeed already been implemented for simple “commercial” computer game demos ².

Even more recently, in the context of BCI use as a secondary input to enhance the interaction with computing systems, the use of *Passive* BCIs has been proposed [39, 203]. In this case the speed of the BCI, to detect a brain state, is no more an issue and moreover the BCI is not meant at all to substitute ordinary interaction devices, but is meant to supply to the system completely different information, which could hardly be obtained with other means, as will be described in Sec. 3.3.

3.1 GENERAL ARCHITECTURE

A VR environment can be defined as an immersive system providing the user with a sense of presence by means of interaction devices with a real-time simulated synthetic world [109]. The user has to be able to interact with the environment in real-time with input devices as keyboards, mice, data gloves, motion trackers, eye

¹These performance increase and shorter learning time, are probably related to the increase of the user’s sensation of presence and immersivity lead by VR environments, enhancing the perceived feedback.

²<http://store.neurosky.com/products/the-adventures-of-neuroboy-bci-technology-demo>

trackers, BCIs etc. While, on the other side, the user has also to be able to receive feedbacks from the system about the virtual world state, thanks to output devices as ordinary displays, large immersive displays, head mounted display, spatial sound systems, haptic devices, etc.

According to [14] typical interaction tasks with VE can be described as belonging to one of the following categories:

- *Object selection*: it consists in selecting an object among the set of proposed ones in the virtual world.
- *Object manipulation*: it consists in changing some of the properties of an object which has commonly been previously selected.
- *Navigation*: it consists in instructing position changes of the user in the virtual world in order to explore it.
- *Application control*: it consists in instructing command to the system to change some of its settings, properties or behaviors.

All of the cited interaction tasks can be performed using a BCI and various examples will be given in Sec. 3.2. It is worth to notice that all of these interaction tasks, being listed in general in [14] for 3D-VE, are meant to encompass primarily explicit interaction tasks, where the user is willing to instruct a command to the system. All of these tasks could indeed be implemented using *Active* or *Reactive* BCIs.

Actually they could be implemented also using *Passive* BCIs, but interesting applications using this BCI category are commonly focused only on the *Application control* task. *Passive* BCIs permits to implement implicit interaction tasks, thus where the user is unaware of the instruction and execution of commands by the system. The user would commonly realize the system is changing, but it would be unaware of being the “source of the signal” provoking the system to change. These kind of BCIs will be discussed separately in Sec. 3.3.

Each BCI modality may be more or less suitable for each of the listed tasks, e.g. Motor Imagery and SSVEP based BCIs, being possibly asynchronous, are more suitable for navigation and object manipulation, while P300 based BCIs may be

more suitable for object selection. Nevertheless anyone of them could be used for each of the listed tasks, selecting the right paradigm [109]. For example in [27] a SSVEP based BCI has been used to implement a BCI Speller.

In general, a BCI based VR setup, is implemented using two distinct software frameworks, one has the duty to acquire brain signals, compute the corresponding features and classify them in order to provide commands to the VR environment, while the second one has the duty to render and visualize the environment. The two software have to be able to communicate with each other using standard protocols. In particular, every BCI based VR setup needs at least the brain signals processing framework to be able to instruct commands to the VR framework, but for synchronous BCIs (e.g. P300 based BCIs) and in some cases also for asynchronous ones (e.g. for the initial classifier training), a bidirectional communication is needed to provide events triggers to the processing framework. If the used BCI modality needs the presentation of stimuli, synchronization messages between the stimuli presentation software and the signal acquisition and processing framework could be needed as well.

The VR environment for BCI based VR setups could be implemented with various software frameworks which are commonly used in this field also for non BCI based environments, from custom OpenGL [172] implementations to more complex systems as Ogre3D³, Panda3D⁴, XVR⁵, etc. These software commonly implement (natively or thanks to contributed libraries) a standard communication protocol to receive inputs and send outputs to generic devices, known as Virtual Reality Peripheral Network (VRPN), which is widely used in the VR field [179].

On the other side, exist also various software frameworks for brain signals acquisition and processing aimed to BCI applications [17]. These kind of software have commonly smaller user communities and have often been developed in the context of single BCI research groups. One that is recently gaining attention, which has been developed in particular for the integration with VR environments, is the OpenVibe framework [163]. This software platform has been recently developed and thanks to a web based community, growing around an on-line support forum⁶,

³<http://www.ogre3d.org/>

⁴<https://www.panda3d.org/>

⁵eXtremeVR 3D software, VRMedia, <http://www.vrmedia.it/>

⁶<http://openvibe.inria.fr/forum/>

is being rapidly expanded in terms of offered functions and features. Indeed it will be adopted also in this work and will be described in detail in Sec. 5.3.

3.2 ACTIVE AND REACTIVE BCIs APPLICATIONS

In this Section will be reviewed existing applications of BCI technologies in the context of VR environments for explicit interaction. They will be divided in accordance to the used BCI modality and thus according to the neurophysiological signal used for the features extraction.

3.2.1 MOTOR IMAGERY BASED

MI based BCIs have been the firsts to be used in the context of VR environments and are definitely based on the modality which was adopted by most of the research groups in this field. This is probably due to the fact that is based on a well studied neurophysiological signal [144], but also, as already mentioned, because of the fact that contrary to SSVEP and P300 based BCIs, MI ones do not require any external stimuli presentation [109]. This lead to simpler systems, not requiring precise synchronization mechanisms and events triggering.

The positive impact of using VR environments as visual feedback for MI based BCIs was initially investigated in [101]. In this work, a two class left/right hand motor imagery BCI has been implemented, showing at first to the user a simple bar feedback moving in the same direction as the detected imagined hand movement (left vs right).

The same experiment using the same signal processing method was repeated using as visual feedback an immersive VR environment provided by two different visualization devices: a CAVE⁷ and an Head Mounted Display (HMD). In these cases the provided feedback was a change in the user position inside the environment, according to the detected mental state (imagined left/right hand movement). No differences between using HMD and CAVE visual feedback have been

⁷A *Cave Automatic Virtual Environment*, better known by the recursive acronym CAVE, is an immersive virtual reality environment where projectors are directed to three, four, five or six of the walls of a room-sized cube. The name is also a reference to the allegory of the Cave in Plato's Republic.

reported in respect to the BCI performance, but all users performed better in these environments compared to the standard 2D bar feedback.

Further experiments confirmed the initial results both for hands imagery and foot imagery [109], thus it is now known as a proven fact that feedback provided by immersive VR environments facilitate the users to obtain higher performances using MI based BCIs.

In [167] is proposed a 3-class self-paced MI based BCI, where left/right imagined hand movements are used for steering, while foot imagined movements are used to walk in a VR environment in order to provide the user all the commands needed for the navigation task. Although this proved to work, it also highlighted some of the limitations given by this approach. Increasing the classes to be discriminated, the detection accuracy decreases, but moreover, the application revealed to be very tiring for the users, since to walk between two points in the VR environment they had to continuously perform one of the three mental tasks.

Different approaches have therefore been proposed, trying to use only few classes⁸ and more complex interaction techniques based on objects selection. For example, an user instead of issuing to the environment the direction in which he/she would like to walk, may select a point of interest and the system would automatically “walk there” [109].

One of the strength of MI based BCIs in the context of VR environments is that they can be easily implemented as self-paced BCIs and that they do not need high accuracy synchronization mechanisms with the software render of the environment, since there is no need of stimuli presentation. On the other side, one of the most severe limitations is that they are able to discriminate only among few mental states and indeed they were used mainly for navigation, where at most 3 states are commonly enough.

3.2.2 P300 BASED

In contrast to MI based BCIs, P300 based BCIs need visual stimuli presentation and were used mostly to accomplish object selection tasks.

⁸As two classes left/right hands imagined movements or one class “BCI switch” relaying on foot imagined movements.

As introduced in Sec. 2.3.1, P300 based BCIs rely on the detection of the P300 waveform, which is elicited after the presentation of a stimulus which is of interest for the subject, among the presentation of other similar stimuli which are not of interest for the subject. Consequently, the same P300 waveform classification mechanism can be used to select among a virtually infinite number of targets, providing a BCI very well suited for object selections. Despite of this, in practice, the number of usable targets is not infinite because of the time needed to show all of them to the subject, anyhow using paradigms as the one presented in Sec. 2.3.1 for the P300 speller, BCIs with tens of targets could be implemented keeping the selection time reasonable. Interestingly, the P300 response is even more pronounced (and thus easier to detect), if more targets are presented, since its strength increases as the likelihood of the presentation of the searched target decreases [93].

One of the first implementation combining VR and P300 based BCIs was presented in [7], where in a simple virtual smart home an user could control different appliances as TV set or lights using the P300 BCI modality. The user had simply to gaze at the appliance he/she wanted to turn on/off, while 3D spheres were randomly appearing over the objects. Users, were simply asked to count the number of spheres appearing over the object of interest. Only when a sphere was shown on the object of interest a P300 waveform was likely to be elicited and thus detected after few presentations, giving to the system the information regarding which object had to be turn on or off.

More recently more complex implementations have been reported, with similar goals towards smart home control, as the one proposed in [62]. In this implementation a whole virtual house with 6 different rooms was designed, containing different appliances in each room, reaching 200 control commands. Interestingly, in this implementation the stimuli presentation was not integrated in the environment, but was showed in a separated monitor and an inertial head tracking device was used to discriminate if the users were gazing at the control monitor or at the 3D environment. This lead to the possibility to turn on and off the BCI control according to the will of the user to instruct a command, avoiding their distractions to cause misclassifications, which in the context of synchronous BCIs, as the P300 based ones, is not trivial [3]. Seven different control masks (i.e. icons matrices) were selectable from the user in the control monitor, containing different icons enabling

him/her to select predefined positions where to move or predefined commands for the appliances. Consequently, in this application also the navigation task was implemented, but it is worth to notice that in contrast to the navigation provided by MI based BCIs, in this case it was not left to the free will of the user, since only some point of interest were presented to let the user choose among them. This is sometimes referred as a goal oriented BCI control approach [109].

These experiments proved the feasibility of using P300 based BCIs in the context of objects selection and manipulation tasks in VR environments, highlighting the fact that graphical icons could be used instead of letters (as in the classic P300 speller), but also the fact that the same classifier trained to detect a symbol could be used for different symbols without retraining [109]. Although a per-subject initial training phase is needed anyway.

On the other side these experiments, in particular the ones presented in [62] and [3], highlighted one of the weakness of the P300 modality which is its synchronous nature, requiring particular expedients to implement an asynchronous control over the system.

Apart from the technical problem given by the synchronous nature of the P300 modality, the work presented in [62] highlight another issue of this approach. The BCI control has been compared with a gaze-based selection method coupled with wand navigation with respect to the user perceived sense of presence in the VR environment. Results suggest that the P300 BCI implemented with a second monitor for command issuing, gives lower presence scores than the gaze-based approach, probably due to the fact that presence is often “break” by the gaze shift toward the control monitor [109].

Another interesting approach proposed to overcome this weakness is the use of hybrid BCIs integrating both P300 and SSVEP modalities in the same system in order to use the SSVEP detection as a switch to communicate the will to perform a selection, to be later made using the P300 modality [141]. The same approach has been successfully applied also in the context of VR environments again for a smart home control application in [46].

3.2.3 SSVEP BASED

As for the P300 BCI modality, also for the SSVEP BCI modality the presentation of visual stimuli is needed. As previously mentioned in Sec. 2.3.3, a SSVEP response at a given frequency (and its harmonics) is elicited mainly in the occipital region of the cerebral cortex, whenever a subject attend a repetitive change in the parameters of a visual stimulus, at a constant frequency. Interestingly, SSVEP responses can be modulated by the user attention in respect to the stimulus, which means that the SSVEP response will be stronger when the user focuses his/her attention on the stimulus.

As is detailed in Chap. 4, the changing parameter in the visual stimulus could be the color, the pattern, the position, the stereoscopic depth, etc. Anyhow, for BCI applications the most used parameters are color and pattern. A common way to provide a stimulus with a repetitive change in color for SSVEP elicitation is to show to the subject a light flickering at a fixed frequency, while, on the other side for pattern changing stimuli, is often use a checkerboard where black and white checks invert their color at a fixed frequency. This topic is addressed in detail in Sec. 4.1.

The main issue to face in the implementation of a SSVEP based BCI control in a VR environment, as for the P300 based BCI, is therefore to find a way to embed the stimulus presentation in the environment. A trivial way could be to attach flickering lights around the displayed environment as shown in Fig. 4.1.4, but this would lead to a similar weakening of the sense of presence of the user as the one discovered in [62] for the P300 modality.

Since for SSVEP based BCIs is much easier to implement asynchronous operations, a more interesting approach would be to embed the stimuli presentation in the VR environment itself and indeed a first implementation of this kind have been presented in [95]. In this work a simple two classes SSVEP based BCI is used to control a character in a 3D gaming environment, where the user focusing the attention over one of the two available flickering stimuli could control the balance of the character engaged in a tightrope walking task. A screenshot of the environment is shown in Fig. 3.2.1.

In a later work [184] the SSVEP modality has been used also in a more immer-



Figure 3.2.1: A screenshot of the first 3D game controlled by a SSVEP based BCI embedding the stimuli presentation in the environment. Figure taken from [95].

sive application where the VE was displayed in a CAVE system and where the BCI control was used to implement the navigation task. Also in this case a two-class BCI was used to discriminate between left and right steering.

The integration of the SSVEP stimuli presentation in VEs, in contrast to the use of external stimuli generators as flashing lights, poses various limitations; first of all the allowed stimuli frequencies are limited by the screen refresh frequency and secondly, software tools are needed to provide an accurate stimuli synchronization with it in order to present stable frequency flickers, as will be addressed in detail in Sec. 4.1.3. Furthermore, another limitation is given by the fact that an aesthetically pleasant solution granting the users' sense of presence in the VE has to be proposed, thus a balance between effectiveness of the stimuli and "natural" objects granting the user presence has to be reached ⁹.

The implementations mentioned so far, although embedding the stimuli presentation into the VE, privileged the effectiveness of the presented stimuli to elicit SSVEP responses in the users. Indeed they rely on flickering squares or checkerboards overlaid over the screen, as shown in Fig. 3.2.1, causing anyway a lowering in the users' sense of presence. Later works, in order to address this issue, moved

⁹Different stimuli shapes and colors have been proven to elicit SSVEP with different intensities [188]; this topic is addressed in detail in Chap. 4.

towards stimuli more tightly integrated in the environment.

In one of the different scenarios proposed in [49], flickering stimuli were fixed to the hands of an avatar and hence were dynamically following every avatar movement; the user gazing at one of the two of them could instruct the avatar to push one of the two buttons in front of his hands.

An even more natural and ecological approach, named *nimesis interface* has been followed in [105] where flickering stimuli have been integrated in the wings of some butterflies depicted in the VE scene. The user, gazing at one out of three butterflies (one on the left, one in the center and one on the right) could navigate the VE. Presented results suggests that the usage of a controller integrated within the virtual scene along with the feedback seems to improve subjective preference and feeling of presence, despite of reducing performance in terms of speed. The authors indeed suggest that flickering stimuli presented in an ecological way should be used for controls, in systems where performance demands could be relaxed, in benefit of an improvement in the interaction naturalness.

Further implementations in this direction have also been presented as the one proposed in [104], which moreover introduce the use of a goal-driven paradigm, to compensate the loss in performance given by the more-pleasant/less-effective flickering stimuli with a more sophisticated interaction mechanism, implementing a sort of shared control between the user and the system. The approach of shared control is indeed gaining attention in order to move towards the use of BCIs to instruct higher-level commands, while adopting Artificial Intelligence (AI) techniques to translate them in lower-level controls [104, 146].

The interest in a more pleasant environment for the users, reducing eye fatigue and increasing the sense of presence (although to the detriment of performance), is moving the research in this field also towards higher flickering frequencies, which should be privileged since are less annoying for the user and less consciously perceptible [10].

Moving beyond traditional VE towards the overlaying of SSVEP eliciting flickering stimuli over the real world, in [48] has been presented an Augmented Reality (AR) environment controlled by a SSVEP based BCI. In a pilot study, two out of three healthy volunteers successfully performed a navigation task using an HMD where flickering stimuli were superimposed over the real world view acquired by

a camera mounted over the HMD.

This is an extremely interesting result in view of SSVEP based BCIs utilizable also in the real world, without the need of particular hardware to implement flickering stimuli over real objects in the environment.

3.3 PASSIVE BCIs AND HUMAN-MACHINE SYSTEMS

All of the works mentioned so far implement what is known as explicit interaction and thus an interaction where the user is able to consciously instruct a command to the system. On the other side, in this section implicit interaction by means of BCIs will be addressed and the concepts of Human-Machine Systems (HMS) and Affective BCI (aBCI) will be introduced. In particular will be addressed the prospective use of *Passive* BCIs for environments/games adaptation with respect to the concept of flow.

The research field of HMS investigates the interaction between a technical system and its user in general [199, 202]. The HMS differs from the more general and well known field of Human-Computer Interaction (HCI) in that it focuses on complex, dynamic control systems that often are partially automated and may exhibit adaptive behaviors with respect to the user. Automated adaptation aims at designing technical systems which can interpret the current state of the user and change the properties of the interaction, or others environment parameters, according to it.

There exists a whole field of study grown around the term *Affective Computing*, with the main purpose of the automatic recognition of affective user states [145], and great efforts have led to promising results for user state estimation via behavioral and physiological signals.

User affective state may be inferred by visual and audible behavior e.g. by video recording and audio recording the user while interacting with the system, or alternatively by recording physiological responses as heart rate, respiration, Galvanic Skin Response (GSR), to derive the user's affective state [146]. Behavioral observations demonstrated to be a good signal source to infer the user state, but pose different issues related to an high inter-subject variability, potential social context biases, but also to the ease for the user to mimic fake behaviors deceiving the recog-

nition system [123]. Physiological signals in this context seems to be more reliable and in particular EEG signals, with respect to peripheral signals (e.g. GSR, heart rate, breathing rate, etc.), would be the hardest to be faked, or to be modulated by environmental conditions [199].

A system able to detect the user state from physiological measurements and able to adapt to it, would find various applications in different research fields, according to the different detectable states, from industrial applications to the improvement of user experience [199].

One of the more appealing applications in the context of Computer Games, Serious Games and learning environments in general, which is recently gaining attention, is the adaptation of games difficulty according to affective game-related user states [109, 146]. This kind of application poses its bases on the concept of *flow* that will be introduced in Sec. 3.3.1 and some works towards this direction, together with other *Passive* BCI applications in the context of VE and Computer Games will be reviewed in Sec. 3.3.2.

3.3.1 THE CONCEPT OF FLOW

The study of the concept of *flow* started in the seventies, by Csikszentmihalyi and other researchers working in the wider field of *creative processes* [38]. The main observation leading to this study regarded artists working on paintings. It was noticed that when working on their art creations they were completely single-mindedly concentrated on their work, disregarding hunger, fatigue and discomfort, while rapidly losing interest about their paintings whenever they were completed [132].

Flow research aimed to study this phenomenon of intrinsically motivated, or *autotelic*, activity, able to produce a rewarding sensation apart from its end product or any extrinsic good that could result from the activity itself.

Trying to give a brief definition, the *flow* could be defined as the state of mind which makes us to stay focused on an activity. Where with the term “stay focused”, is meant a state in which all of our attention is directed toward the activity itself. The *flow* concept applies to any human activity and there are several subjective and objective factors able to foster or discourage the arising of the *flow* state in a person. In general, the conditions allowing the *flow* state are mainly [132]:

- the presentation of challenges or opportunities that stretch existing skills, without overmatching nor underutilizing them and thus a sense of engaging challenges appropriate to one's capacities;
- the presentation of clear goals reachable with one's skills which would lead to an immediate feedback about the progress being made.

People experiencing the state of *flow* often describe an intrinsically rewarding sensation, leading them to the will to continue to perform the activity which provoked in them this sensation, such that the end goal is often just an excuse for the process.

As previously mentioned, one of the conditions allowing this state in a person is the presentation of a challenge that is considered to be tackled with his/her own skills. Consequently a challenge which is considered too hard to be tackled, or a challenge which is considered too easy, will prevent to get into the *flow* state, but on the other side, will probably lead a person respectively into two other unpleasant states named *anxiety* and *boredom*.

Flow activities must manage to keep the user in the narrow margin of challenge that lies between boredom and frustration, since both of these unpleasant extremes cause our mind to change its focus to a new activity. Csikszentmihalyi called this margin the *flow channel*, showed in Fig. 3.3.1.

In reference to Fig. 3.3.1, when a person faces a workable challenge and succeed to enter the *flow* state (A_1), after some time he/she will probably increase his/her skills, learning how to face it. Consequently, the same challenge will hardly keep a person in the *flow* state if presented again, leading him/her in the unpleasant *boredom* state (A_2). On the other side, if the challenge difficulty increases faster than the person's skills, he/she will enter in the unpleasant *anxiety* state (A_3) feeling the frustrating sensation of being unable to face the challenge. Once out of the *flow* state, in (A_2) or in (A_3), a person would like to enter it again (A_4), but if no challenges of the right difficulty are presented in the current activity, he/she will soon completely lose interest in it, looking for other activities able to lead again to the *flow* state.

The reason why humans experience the state of *flow* is surely related to its ben-

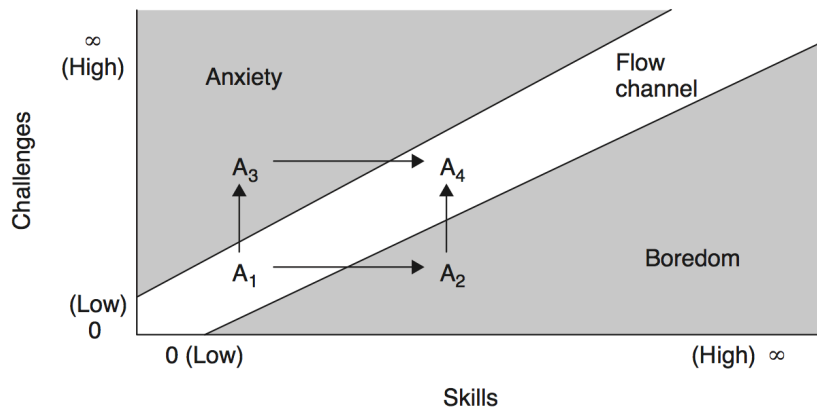


Figure 3.3.1: Simple schematic representation of the flow channel as a three user states plot. Figure taken from [166].

efits with respect to learning processes. The ability to enter in this intrinsically rewarding state has probably been hard coded in the human brain by natural selection, in order to provide as a reward a pleasant sensation, while learning new skills. More the subject try and succeed to enter the pleasant state of flow, more skills he/she will learn, which from the natural selection point of view is a great advantage.

Mammalians in general, but humans in particular, developed the behavior of “playing games”, characterizing mainly infancy and indeed games actually are nothing else than intrinsically rewarding activities that may lead the subject/user the possibility to enter the state of *flow* and thus to enjoy the learning of new skills. One of the reasons that is believed to be at the basis of the “success” of the Homo Sapiens specie and later of the Homo Sapiens Sapiens sub-specie, is actually that it has been evolving towards a growing *psychological-neoteny*. *Neoteny* can be roughly defined as the retention by adults of traits previously seen only in juveniles [120], it is a subject studied in the field of developmental biology and many prominent evolutionary theorists propose that it has been a key feature in human evolution.

Psychological-neoteny lead in humans, in contrast to other mammalians, the ability to enter the state of *flow* also in adulthood¹⁰ and thus to enjoy various kind of

¹⁰Assuming that a generic young mammalian while playing could enter in a state similar to the human state of *flow*.

activities as games, hobbies, arts, etc. which are performed just for the “fun” of performing them. The strong link between *psychological-neoteny*, gaming and learning is indeed demonstrated by various theories, one of which for example is that highly educated people and eminent scientists usually demonstrate more neotenous psychological traits [32].

The *flow* theory has been consequently deeply studied in the fields of game design, serious game design and learning processes, since one of the most important features of a gaming or learning environment is to be able to keep the players playing and thus into the state of *flow* [166].

The same holds true in particular for computer based environments where the difficulty could be much more easily controlled than for real games as tennis, soccer, etc. In [170] the author indeed asserts: «Some may comment that Csikszentmihalyi seemed to have video games in mind when he developed the concept of *flow*», moreover: «[video] games possess ideal characteristics to create and maintain flow experiences in that flow experience of video games is brought on when the skills of the player match the difficulty of the game».

The reason why different levels are commonly present in Computer Games is precisely in order to provide an increasing challenges difficulty in order to keep the player in the *flow* state, as shown in Fig. 3.3.2.

Actually, the ideal challenges, according to [166], to provide the best gaming experiences, should follow the line depicted in Fig. 3.3.3 swinging between the *anxiety* and *boredom* lines without crossing them. In this fashion the user is provided with alternating sensations of easily reachable rewards and challenging tasks.

According to a more complex model of the *flow channel* introduced by Csikszentmihalyi in 1997, adding more possible states, as shown in Fig. 3.3.4, the swinging line in Fig. 3.3.3 would translate in a swinging between the *control* state, in which the user feels to sensation to have everything under control and the *arousal* state, in which the user feels to be challenged by a new dare.

Unfortunately the borders of the *flow channel* are extremely subjective, they depends on the initial skills of the user, on the initial interest elicited by the activity, but also on its learning speed. One of the main works of a game designer is there-

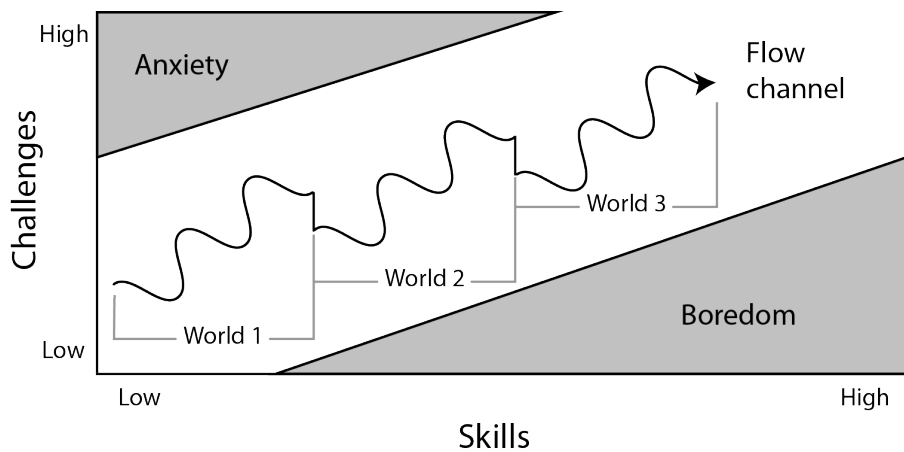


Figure 3.3.2: Simple schematic representation of the flow channel as a three user states plot, highlighting how Computer Game levels or “worlds” are designed in order to keep most of the users inside it. Figure taken from <http://indiedevstories.com/2011/08/10/game-theory-applied-the-flow-channel/>, adapted from [166].

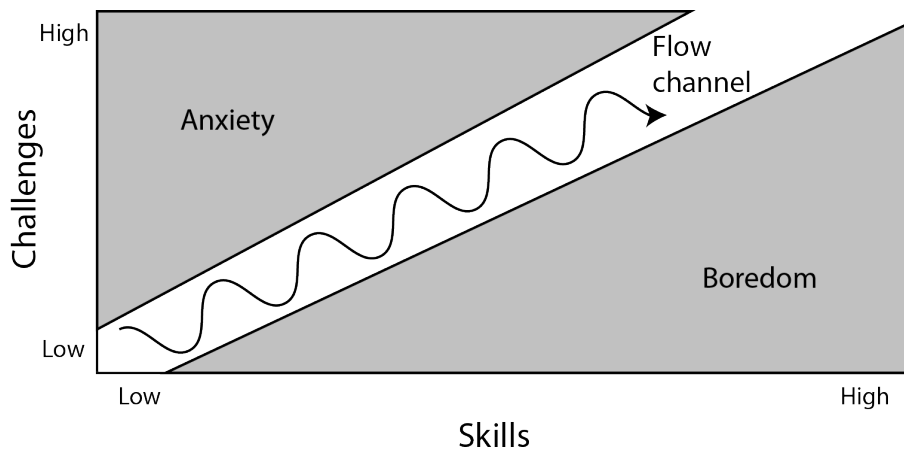


Figure 3.3.3: Simple schematic representation of the flow channel as a three user states plot, highlighting the ideal path the user should follow for a maximal engagement. Figure taken from [166].

fore to fit the game into the flow channel of the maximum number of potential players. The main problem in doing that, is given by the absence of a feedback from the user, telling the game (or system in general) if the user is close to one of the two borders of the *flow channel*. Thus the game difficulty change has to be previously programmed and is therefore implemented as an *open-loop* system, apart

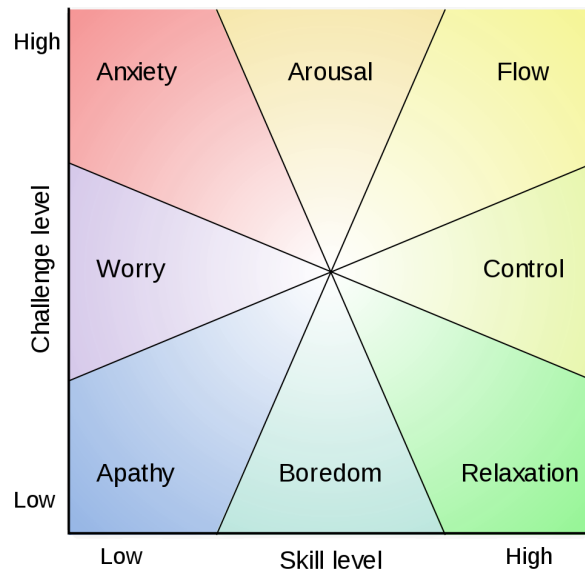


Figure 3.3.4: A more complex model of the *flow channel* introduced by Csikszentmihalyi in 1997. Figure taken from <http://en.wikipedia.org/>.

from the “levels” or “worlds” divisions.

Anyhow, the game difficulty adaptation would be an easily manageable task if an *anxiety* and a *boredom* detector were available, or at least a *flow* state detector. If this would be the case, a *feedback-loop* could be implemented in order to change a generic game or learning environment difficulty according to the state of the particular user playing with it [146].

As already mentioned, some research works succeeded to link features extracted from physiological signals, and EEG in particular, to users’ mental states elicited through controlled environments, among which also the *flow*, or its related game states [9, 31, 129, 148].

Some proof of concepts of closed-loop implementations have been provided too, using multiple physiological signals as the one presented in [153], although the on-line experiment has been tested only on two subjects and further tests seems not to have followed.

Anyhow, although game play has been intensively studied, the underlying neurobiology is still poorly understood and is currently still under investigation [88, 196].

A key concept linked to the flow state seems to be the widely studied cognitive process of attention, which, although not having a unique definition, is largely conceptualized as the means by which the brain chooses information for further processing [196]. The flow state is indeed defined also by Csikszentmihalyi, as a state in which the attention of the subject is highly focused over the task eliciting the state, thus concentrating most of the available brain resources to the particular task. It is known in fact that when a subject is experiencing the state of flow he/she tends to ignore external stimuli, but also internal ones as hunger, fatigue, etc.

The psychophysiological construct of attention can be seen from two point of views; attention as an arousal mechanism, identifying the state of physiological reactivity of the subject to external stimuli, and attention as a selective process, identifying the focusing of brain processing resources on a particular stimulus or object.

In particular, when speaking about attention in the rest of this work, I will refer to attention as a selective process and when speaking about visual attention I will refer to its definition as a two-stage process: in the first stage, attention is distributed uniformly over the external visual scene and processing of information is performed in parallel, while in the second stage, attention is concentrated to a specific area of the visual scene (i.e. it is focused), and processing is performed in a serial fashion.

Another key concept linked to the flow state is related to its pleasantness and gratification for the individual experiencing it, which on its turn has to be related with the pleasure and reward networks of the brain.

In fact, the theory proposed in [196], aiming to explain the neural processes characterizing the flow state, propose (and in part demonstrate using fMRI scans of subjects playing a computer game), the appearance of a neural synchronization between attentional and reward networks of the brain ¹¹.

The link between the flow state and how the user attention is focused, is indeed a key concept that in this work has been proposed to be exploited for prospective *Passive* BCI in the field of game difficulty adaptation.

¹¹Synchronization of neural networks is known to be an index of the co-interaction of different brain areas and of the exchange of information between them [137, 187].

3.3.2 PASSIVE BCIs APPLICATIONS

The recently introduced concept of *Passive* BCI is gaining attention in the field of HCI and in particular in the field of HMS, in order to implement implicit interaction between the user and a generic system [199]; this is therefore in particular very interesting for Gaming applications and the interaction with VR environments [99].

As introduced in Sec. 2.4, *Passive* BCIs are characterized by the fact that the user is unaware about the BCI itself; the user is requested just to interact with the system using ordinary interaction devices, while the system thanks to the features extracted from the user neural activity can “implicitly” adapt to the user state.

The concept of *implicit interaction* has been used with slightly different definitions, in different research fields, but as stated in [60], they all seems to refer to the same idea: “An interaction process that is not based on direct, explicit or voluntary action of the user, but more on the state of the user in a particular context. Both the user’s state and the given context can thus be associated with the expression *implicit information*”. As already mentioned, behavioral and physiological data could be used to acquire implicit information, but, recalling the general definition of BCI, only neuro-physiological data is relevant to the *Passive* BCI research field.

The main limitations to possible practical applications of *Passive* BCIs are given by which mental states could be detected, by the detection accuracy and the ease of use (e.g. long calibration sessions for classifiers training limit practical use). In [60] is given an overview of the detectable states reported in the literature which could be adopted for prospective *Passive* BCIs, with the respective EEG features utilizable. Most of the research works presented seems to fall in one of these four categories:

- *Relaxed alertness estimation*: The EEG alpha band has been related to the idleness of cortical areas since long time ago; therefore its power can be used to estimate an index related to the user relaxation or alertness. Despite of this apparently simple observation, very complex behaviors have been observed and studies on various components within this band could be considered as a per se research field [169].
- *Mental workload estimation*: The analysis of workload from EEG data has a

tradition in the psychological community. Most of the works in the past concentrated on the EEG clinical bands power changes ¹², trying to identify the most significant features, commonly derived as bands power ratios, linked to specific brain states. Despite of this, there is still no consensus on the effects of workload on the EEG signals. Moreover, terms as mental workload, task demand, engagement, vigilance, and others are often used interchangeably in literature to describe a human internal state of mental effort [71].

- *Mood and emotions assessment*: Also the detection of user emotions and mood has a quite long tradition in EEG signal analysis [57] and a huge amount of research works have been produced in this direction using as features band powers asymmetries [69], ERP waveforms [68], phase synchronization [36], etc. Despite of this, also in this case there is still no consensus on the best features representing emotional states and moreover different models and categorization of human emotions exist.
- *Perceived error detection*: The detection of the so called Error-related potential (ErrP) is probably one of the most interesting techniques used lately to implement effective *Passive* BCIs [202]. ErrPs are a reaction to an error committed by the subject himself or by the system trying to interpret her/his intentions, thus they can be used by a generic system to implicitly obtain the information that according to the user something “wrong” happened and thus this information could be used to change/correct its behavior. ErrPs are supposed to be generated in the anterior cingulate cortex (ACC), which is crucial for regulating emotional responses [57].

Although presented far before the term *Passive* BCI was coined, a first pilot study of this kind of BCI is presented in [150], with the aim to estimate the user “engagement” in a task to use the information in a closed control loop for system adaptation. Although the term “engagement” was used, the work is mainly concentrated on the mental workload estimation and as the best feature has been identified the band power ratio $\beta/(\alpha + \theta)$.

¹²EEG clinical band powers are the ones described in Sec. 2.1.4.

Concerning more recent implementations in particular in the context of Computer Games, some proof of concepts of *Passive* BCIs exist, although as far as I know the implicit information has never been used for game difficulty adaptation yet. In [67] and [122] for example, the alpha power activity has been used to change the game behavior according to the user relaxation in order to alter the avatar graphical aspect or the avatar controllability, respectively. In another Computer Game using an implicit BCI presented in [200] has been exploited the ErrP EEG component, where the user has to accomplish a task while the system is able to adapt its behavior according to the possible detection of an ErrP in the user EEG denoting its feeling of loss of control. In a very recent work [99] is given also a review regarding applications to VR environments and in particular the use of a passive mental workload estimation is used to control an haptic feedback.

Despite of the fact that, as already mentioned, several studies showed that a classification of several affective user states is in principle possible, using neurophysiological signals in general, and EEG in particular, few effective practical implementations of closed-loop systems exist yet [146].

Concerning in particular the game related states, introduced in Sec. 3.3.1, various studies highlighted how their induction in the users using specific gaming environments, was able to modify band power features extracted by the users' EEG signals [9, 129]. The reason why is still missing a practical *Passive* BCI implementation able to control the game difficulty, starting from these studies, is in my opinion related to the fact that the underlying neurobiology of the states willing to be detected is still not known enough [88, 196].

As could be glimpsed from Chap. 2, the functioning of the brain is quite complex and the variation of EEG signals powers across the cerebral cortex is just a shadow of a huge number of extremely complex functions concurrently happening. The power in a specific band of an EEG signal is therefore likely to be the non-linear sum of the contributions of different brain processes and thus is an huge challenge to understand the reason why an increase or a decrease of the power happened.

From several studies it is known that game difficulty modulates different band powers, but it is hard to know if the modulation happened actually due to the elicitation of a particular affective state or due to other factors, e.g. the different stimuli entering the visual system. One of the main problems is then given by the fact that

the used modality to elicit the investigated states in the subject, commonly affect the neuro-physiological signals as well and thus a general consensus on a set of signals features strongly correlated with a particular flow state is still missing.

In [146], speaking about emotions detection, where a similar problem is faced, is proposed the use of multi-modal stimuli for emotions elicitation [121], in order to search for features tightly related to the elicited emotion and not to the stimulus used to elicit the emotion itself.

The main problem could be reduced to the issue of finding the right feature, or set of features. While a comprehensive understanding of the phenomenon is still missing, from the neuroscience point of view, a possibility could be in my opinion to concentrate on other better known phenomena which may be modulated by affective game related states.

Using a well known brain response could help the recognition since it is already known what to look for in the EEG signal and moreover technical tools already exists to filter the response of interest from the underlying uncorrelated brain activity.

As proposed in the discussion section in [129], a possibility could be also to analyze the EEG response to certain in-game events, looking for a response modulation given by the game related states.

As pointed out in [88, 196], but also as already noticed from other observations in [38, 132], the flow state is highly connected to how the attention is concentrated over a particular task and how much hard is to divert it. On the other side, as already mentioned, but also as will be detailed in Chap. 4, also the SSVEP response is highly connected to the user attention with respect to the SSVEP eliciting stimulus.

Following this line of reasoning, in Chap. 6, some experiments will be presented attempting to exploit the SSVEP response for prospective *Passive* BCI applications, thanks to its properties that will be detailed in Chap. 4. In particular, the experiment detailed in Sec. 6.3.2 aims to highlight the possible modulation of the SSVEP response with respect to game related states.

4

Steady State Visual Evoked Potentials

An Evoked Potential (EP), in the context of EEG signals, is an electrical potential elicited by the presentation of a stimulus that can be recorded from the nervous system. In particular, in the case of non-invasive EEG recordings, it can be acquired from an electrode positioned on the surface of the scalp. Visual Evoked Potentials (VEP), are EP elicited by a visual stimulation [138].

The main issue related to EP and VEP detection, is given by their low amplitude (in the order of some microvolts) with respect to the spontaneous ongoing brain activity (in the order of tens of microvolts). Since EP and VEP are time-locked to the stimulus that evoked them, a common technique used for their detection (the same used for ERP), is to present several times the same stimulus to later average the recordings time-locking them with respect to the stimulus presentation time as shown in Fig. 2.3.2. On average, the spontaneous brain activity should elide, since it is not time-locked to the stimulus, while the EPs should sum, emerging from the background [138] as shown in Fig. 2.3.5.

Steady-state VEP (SSVEP), as introduced in Sec. 2.3.3, are a particular case

of VEP, where the same stimulus is repetitively presented at a frequency at least higher than 3.5Hz, but more commonly higher than 6Hz. This kind of stimulus is commonly referred as Repetitive Visual Stimulus (RVS). In this case a new stimulus is presented before the transient VEP response to the previous presentation could finish [25], eliciting a steady-state characterized by a periodic nearly sinusoidal response called SSVEP that can be observed in the recorded scalp EEG signal, particularly in the occipital brain region, where the visual cortex resides [143, 188]. A typical SSVEP response is shown in Fig. 4.0.1 in the time domain, obtained using the time-locking averaging technique, and in Fig. 4.0.2 in the frequency domain, using a power spectrum estimation technique.

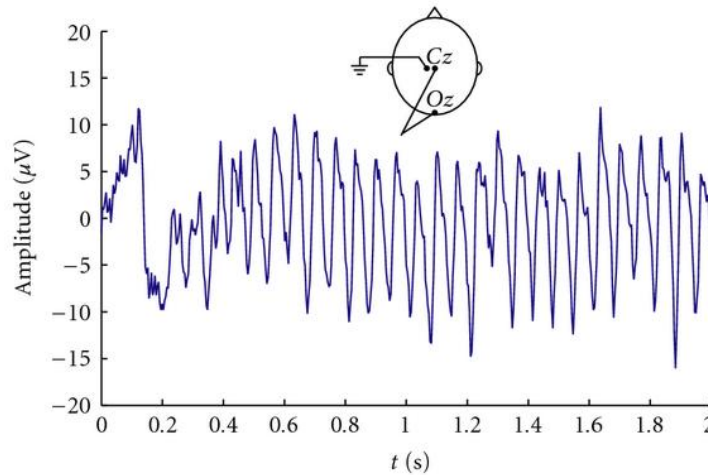


Figure 4.0.1: Waveform of an EEG signal acquired during visual light stimulation with a frequency of 15 Hz as the difference between the signals acquired from Cz and Oz locations (*bipolar derivation*). The SSVEP waveform depicted is the result of a time-locked average of 10 realizations. A transient VEP can be observed at the moment where the stimulation began and a clear oscillation (the steady-state VEP) can be seen afterward. Figure taken from [207].

Although some researchers simplify the SSVEP phenomenon as being nothing more than a sequence of VEPs elicited by each of the visual scene state change, lot of research is operating under the assumption that is safer to assume a less lin-

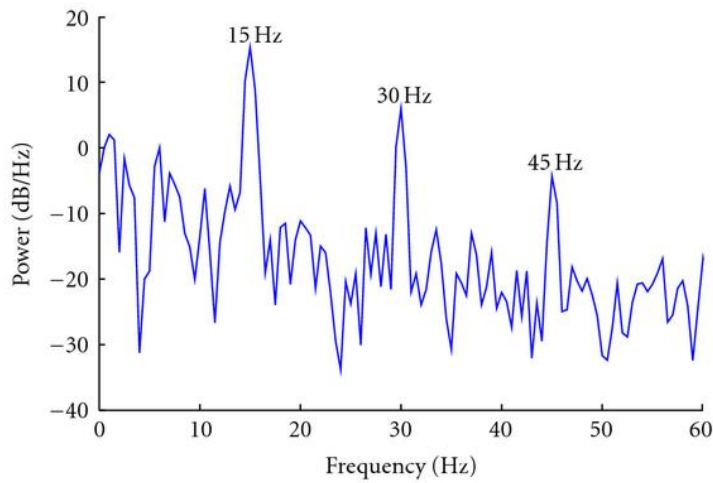


Figure 4.0.2: Frequency spectrum of the same recording shown in Fig. 4.0.1. Figure taken from [207].

ear relationship between the stimulation and the SSVEP response, as discussed in [158].

The stimulus presented to elicit a SSVEP response could be any repetitively changing visual stimulus (e.g. changing color, shape, position, etc.) and accordingly to the “change frequency”, in the EEG signal acquired from the scalp of an user attending to the stimulus, an increase in the power of the signal at the corresponding frequency can be detected. Consequently in the case of SSVEP, summing and averaging different signal epochs corresponding to different presentations of the same stimulus is not needed, since the presence of the response can be detected analyzing the power spectrum in the frequency domain of the recorded signal.

Stimuli are commonly presented by Light Emitting Diodes (LED), or by shapes on a regular computer monitor [207], flickering at frequencies ranging between 6Hz and 40Hz, although it has been proven that SSVEP can be detected also for higher frequencies [151], till 100 Hz. Even higher frequency SSVEP responses could be detected using invasive electrodes, but the meninges, the skull, and the scalp of the subjects acting as low-pass filters, prevent high frequency signals from being detected by surface electrodes.

In this work the SSVEP modality has been chosen because of its high level of de-

tection accuracy [207], the short (or null) calibration time needed, the low number of EEG electrodes required and also for the low BCI illiteracy¹ showed [66], granting high usability for most of the users, also in out-of-laboratory environments.

The use of the SSVEP response for the implementation of BCIs, in contrast to its use for clinical applications or neuroscience research [160], poses new challenges that can be summarized in three needs: the need for an high ITR, the need for comfortable stimuli and the need for practical acquisition devices.

In other words, the need for an high ITR is given by the fact that BCIs are commonly used for interactive applications, thus the time needed to detect a SSVEP response in the EEG signals is crucial for their usability, moreover to implement multiple commands, multiple frequencies should be used and more are the frequencies, more are the available commands, thus high accuracy in the identification of the SSVEP response between the various frequencies is needed to avoid false detections. The ITR, as mentioned in Sec. 2.4.6, is just a measure incorporating the reaction times, the accuracy and the number of available commands for a generic BCI.

The kind of flickering stimuli used to elicit the SSVEP response and how they are presented to the user is highly important to determine a stronger or a weaker response leading to its easier or harder detection in the EEG recording. Stability of the flickering frequency, stimuli shapes, colors, spatial frequency as their affective content for the subject, are all properties known to modulate the SSVEP response which have to be taken into account and that could be exploited for the implementation of SSVEP based BCIs as will be discussed deeply in Sec. 4.2.

The SSVEP response have been studied extensively in the field of vision research starting from the sixties [154, 157, 160], but its use continued in the fields of cognitive neuroscience and clinical neuroscience until today [188] as a tool to investigate the Human Visual System (HVS) functioning, to diagnose some of its possible malfunctioning, but also to study cognitive and affective processing of the brain. In the last years, as introduced in Sec. 2.3.3 it has been adopted widely for

¹The term “illiteracy” has often been used to identify the disability of some users to use a BCI, sometimes also the term “apraxia” is used to express the same concept, but there is still not a consensus in the community on a standard term or definition [2].

the implementation of *Reactive* BCIs [110, 142, 188] and therefore new methods have been investigated to reliably and quickly detect SSVEP responses from EEG recordings, starting from earlier research results.

Research towards better SSVEP based BCIs is moving in different directions that could be summarized in two main groups; researches moving towards BCIs with higher ITR and researches moving towards more natural interfaces using less annoying stimuli to increase the users comfort. Clearly a trade off has to be identified since as will be described also in this Chapter, more comfortable stimuli commonly grants lower ITRs.

In this Chapter will be given a review of the state-of-the-art regarding studies related to SSVEP aimed to the implementation of SSVEP based BCIs, but also studies in other research fields that may have interesting implications for this BCI modality.

Furthermore, in this Chapter will be reviewed and described also the different stimuli presentation techniques and the signals detection and analysis methods.

4.1 STIMULI PRESENTATION

As previously mentioned, in order to elicit VEP and SSVEP responses in the subject's brain activity, a visual stimulation has to be provided.

In the context of SSVEP based BCIs various visual stimuli have been tried by means of different stimulator devices with different performances [207] and sometimes with contrasting results [30, 198]. Anyhow, as a matter of fact, the stimulus properties can strongly influence the VEP waveforms and consequently also the SSVEP response amplitude and frequency distribution.

In this Section will be described the most common classes of stimuli used to elicit VEP and SSVEP responses and the different kind of elicited responses will be compared gleaning information from neuroscience research works and from recent SSVEP based BCIs experiments as well. Moreover, will be given a review on the commonly used devices to present the stimuli to the subjects (or users), highlighting their pros and cons.

4.1.1 CLASSES OF STIMULI FOR VEP EXPERIMENTS

For clinical applications and in particular for VEP elicitation, standards and recommendations exist as guidelines for the stimuli presentation and description to perform VEP experiments.

In [138] two major classes of VEP stimulation are identified: luminance and pattern. Luminance stimulation is usually delivered as a uniform flash of light, while pattern stimulation may be either presented in a pattern-reversal or onset-offset fashion, as will be detailed in the next sections.

PATTERN STIMULUS

The recommended patterned stimulus in [138] is a black and white checkerboard where every checks should be a square and where there should be an equal number of light and dark checks, as the one shown in Fig. 4.1.1.

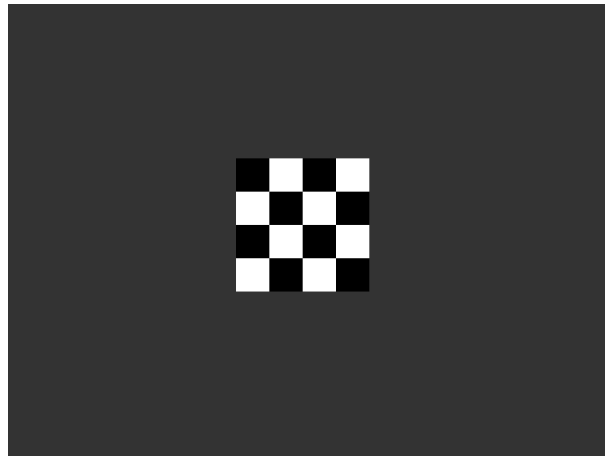


Figure 4.1.1: The checkerboard pattern stimulus

As a standard practice, the dimensions of the checks should be defined as the visual angles subtended by the sides of a single check, since what matter are the dimensions of the stimulus with respect to the visual field of the subject and not its absolute dimensions. Moreover, the visual angle should be measured in degrees and minutes of arc subtended at the subject's eye. The same holds for the definition of the whole checkerboard size, named *stimulus field size* that should be

expressed in degrees of visual angle, with an indication of the field shape, i.e. if it is a rectangular field $\alpha^\circ \times \beta^\circ$ large or a circular field of γ° diameter or radius.

Pattern stimulus luminance should be measured in candelas² per square meter, cd m^2 and the luminance of the white checks should be at least 80 cd m^2 .

The surround of the stimulus should be homogeneously lit, with an average luminance equal to or below the average stimulus luminance. In practice a subdued room lighting with no bright sources visible to the subject has to be used and if a computer monitor is used for presentation, a dark background has to be used.

The location of the fixation point should also be defined in relation to the stimulus field and the fixation point should be positioned at the corner of 4 checks when located at the center of the field.

The *pattern reversal* stimulus consists of black and white checks (as the ones shown in Fig. 4.1.1) that change phase abruptly (i.e., black to white and white to black). There must be no overall change in the luminance of the screen, thus an equal number of light and dark elements has to be displayed.

Otherwise, for *pattern onset/offset*, a pattern is abruptly exchanged with a diffuse background (e.g. the image in Fig. 4.1.1 appears over the background and then disappears). Again the pattern stimulus should be defined in terms of the visual angle of each check. All the previously mentioned recommendations hold, but in this case also the stimulus persistence time has to be taken into consideration. A standard of 100 to 200 ms pattern presentations separated by 400 ms of diffuse background is recommended [138]. It is also specified that the data acquisition system should be set to trigger exactly at the appearance of the stimulus.

FLASH STIMULI

The flash stimulus is defined as the pattern onset/offset stimulus, but in this case there is no pattern; thus on a darker background a patch of solid color uniformly lightened is turned on and then off. In [138] is recommended that VEP should be elicited by a flash that subtends a visual field of at least 20° .

²The candela (symbol: candela) is the SI base unit of luminous intensity; that is, power emitted by a light source in a particular direction, weighted by the luminosity function (a standardized model of the sensitivity of the human eye to different wavelengths, also known as the luminous efficiency function). $1 \text{ cd sr} = 1 \text{ lm}$.

Moreover, the stimulus should be presented in a dimly illuminated room. The strength (time integrated luminance) of the flash stimulus should be measured in photopic candelas seconds per squared meter, cd s m^2 . The background on which the flash is presented should be measured in candelas per squared meter, cd m^2 and the flash should have a stimulus strength from 1.5 to 3 cd s m^2 with a background from 15 to 30 cd m^2 . Furthermore, the stimulus should be presented less than 1.5 times per second ($< 1.5 \text{ Hz}$) in order to elicit VEPs and avoid the elicitation of SSVEP responses.

4.1.2 CLASSES OF STIMULI FOR SSVEP BCIs

To elicit SSVEP responses the used stimuli are commonly the same as the ones used to elicit VEPs, the only difference consists in the presentation rate that has to be faster than “several hertz”, in order to establish the steady-state.

In the context of SSVEP based BCIs the most commonly used stimuli are the *flash* stimuli and the *pattern reversal* stimuli [207]. Despite of this, unfortunately in almost all of the research works most of the recommendations mentioned in 4.1.1 are not taken into consideration and thus the stimuli descriptions is often lacking details.

Apart from these “classical” kind of stimuli, VEP and SSVEP responses, as previously mentioned, could be elicited by any kind of change in the visual field of a subject, thus also different kind of stimuli have been investigated, as shown in Fig. 4.1.2 where most of them are represented.

In Fig. 4.1.2 a further division is illustrated, showing in conjunction to the different kind of stimuli, also the devices which could be used to present them. In the upper box (A) is represented a particular kind of stimulation by means of two different lights positioned on a pair of goggles in order to present flickering stimuli directly in front of the eyes; this could be interesting for particular kind of experiments, but to my knowledge no SSVEP based BCIs exist using this stimulation device. In the second box (B) is represented a common stimulation device for SSVEP based BCIs that is the Light-Emitting Diode (LED) which will be addressed in Sec. 4.1.3; with this device, thanks also to diffusive panels, *flash* stimuli are commonly provided. In the last box (C), different kind of stimuli presentable using

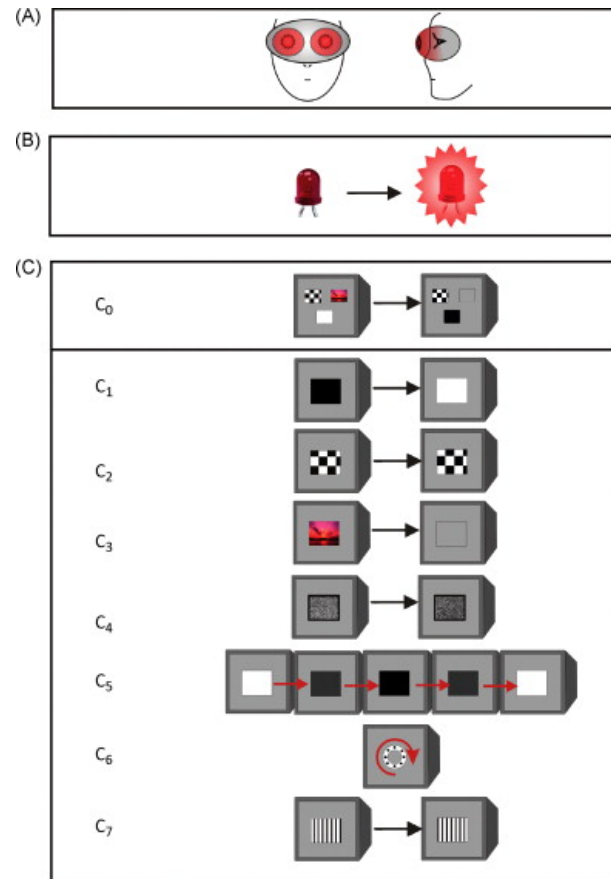


Figure 4.1.2: Representation of different kind of visual stimuli and stimulation devices, able to elicit VEP and SSVEP responses. (A) Flickering light mounted on goggles; (B) Light-Emitting Diode (LED), producing flickering light; and (C) flickering images on a computer screen: (c_0) combination of images that can be used for binocular rivalry paradigms, (c_1) simple square, (c_2) checkerboard, (c_3) image, (c_4) Gaussian field, (c_5) sinusoidally modulated square, (c_6) rotating or moving stimuli, and (c_7) moving vertical or horizontal gratings. Figure taken from [188].

a regular computer monitor are displayed. Most common stimuli, as previously mentioned, are *flash* stimuli (c_1) and *pattern reversal* stimuli (c_2), but also flickering images (c_3) could be used to study the SSVEP response change with respect to the semantic content of the image, as will be introduced in Sec. 4.2.9. Other kind of stimuli ($c_4 - c_7$) were used in the fields of physiology and neuroscience to investigate particular relations between the SSVEP response and other perceptual or cognitive functions of the brain, but were rarely used for SSVEP based BCIs.

4.1.3 STIMULATION DEVICES

To provide visual stimuli to elicit VEP and SSVEP responses, various devices could be used, from specific hardware to consumer display devices.

In the context of SSVEP based BCIs two different devices are commonly used as previously mentioned: Light-Emitting Diodes (LEDs) and computer monitors. LED lights are more suitable for providing *flash stimuli*, while computer monitors could be used for both *flash stimuli* and *pattern stimuli*, although they present several limitations that will be addressed in Sec. 4.1.3.

The SSVEP responses obtainable in the EEG using different stimulation devices have been investigated in various works [207], but conflicting results have been published, e.g. in [198] the SSVEP response obtained using a computer monitor stimulator have been reported to be weaker than using a LED stimulator, but from the results reported in [30] the opposite conclusion is highlighted.

The contrasting results, as suggested in [30], are probably due to the software used and in particular on how the synchronization of the stimuli presentation with the screen refresh is managed, since a not stable flickering frequency may be the source of weaker performances. Fluctuations in the stimuli presentation frequency can therefore result in an unstable EEG frequency spectrum, where will be harder to detect a clear sharp peak at the corresponding frequency and its harmonics.

Another reason for the contrasting results in the literature is that different stimulation methods were often compared using very different stimuli characteristics (e.g. sizes, shapes, colors, duty cycles) and as described in Sec. 4.2, all of these factors contribute in the modulation of the SSVEP response. Moreover in different research works, different data analysis methods were used, thus it is very hard to

compare the results and identify which stimulator could lead the better results.

In my opinion and as stated also in [10], to ask which stimulator device is the best, is actually the wrong question; the only things which matter are the physical characteristics of the provided stimulus. Once identified which is the needed stimulus, a device can be selected from the available ones, according to their ability to display the requested stimulus.

Anyhow, as already mentioned in the case of SSVEP based BCIs, the two most used devices are LED lights and computer displays.

LED LIGHTS

To provide *flash stimuli*, LED lights, or array of LED lights, are preferred with respect to other kind of lamps because of their relatively low latency and fast reaction to onsets and offsets, letting them to be the best choice for a reliable frequency control. They can be driven using a waveform generator or a simple microcontroller programmed with the desired waveform and frequency, as an inexpensive Arduino board [180].

LED lights received great interest from electronic research and industry in recent years because of their high luminous efficacy, reaching for red-orange LEDs, peaks of almost $3 \times 100 \text{ lm W}^{-1}$. Consequently they are widely available on the market, relatively cheap and also power LED exists reaching several watts.

LED lights luminosity, being them from the electronic point of view simply diodes, can be current controlled and few volts of voltage difference are needed to light them. Accurate current control is not too easy to be obtained in an energy efficient way, thus LED lights are commonly dimmed using Pulse Width Modulation (PWM) of their power supply. PWM consist in driving them with a square wave modulated voltage (i.e. turning them on and off very quickly), tuning the duty cycle of the square wave to obtain a current control in average over a time window.

³The lumen (symbol: lm) is the SI derived unit of luminous flux, a measure of the total “amount” of visible light emitted by a source. Luminous flux differs from power (radiant flux) in that luminous flux measurements reflect the varying sensitivity of the human eye to different wavelengths of light, while radiant flux measurements indicate the total power of all electromagnetic waves emitted, independent of the eye’s ability to perceive it. $1 \text{ lm} = 1 \text{ cd sr}$.



Figure 4.1.3: An LED used for SSVEP elicitation mounted in a case covered by a diffusive panel. Figure taken from [10].

This is important to be noticed for SSVEP elicitation, since, although the PWM frequency is commonly in the order of several kilohertz, frequency beating phenomena may arise with the flickering stimulus frequency.

Being the LED commonly very small with respect to the optimal area to elicit VEP and SSVEP responses (several degree of visual angle), they are often used in arrays, or otherwise they are placed behind a diffusive patch as shown in Fig. 4.1.3. To implement SSVEP based BCIs for computer interaction, specific ad-hoc LED based stimulator hardware have been produced and a typical configuration is shown in Fig. 4.1.4.



Figure 4.1.4: Use of LED stimulator devices applied on a regular CRT computer monitor. Figure taken from [44].

When an LED is switched on, electrons are able to recombine with holes within the device, releasing energy in the form of photons. This effect is called electro-

luminescence and the wavelength of the light (corresponding to the energy of the photon) is determined by the energy band gap of the semiconductor. Consequently LEDs are characterized by having a very narrow bandwidth in terms of their emitted light spectrum that is specific to the material used to produce them. Different materials are used to provide LEDs emitting different light wavelengths and thus characterized by different perceived colors.

Wide spectrum LEDs, often called white-LEDs can be obtained using multiple LEDs of different wavelengths (e.g red, green and blue), or by coating near ultra-violet emitting LEDs with a mixture of phosphors ⁴ (e.g. high efficiency europium based red and blue emitting phosphors plus green emitting copper and aluminium doped zinc sulfide). The coating made by different materials enrich the emitted wavelength spectrum with different peaks eliciting in the observer the perception of a white light. Two spectra of two generic white LEDs implementing the two different approaches are showed in the left and right parts of Fig. 4.1.5.

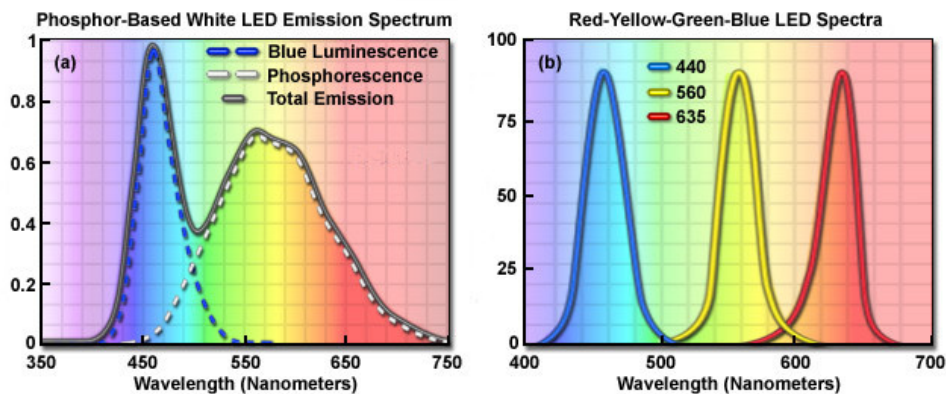


Figure 4.1.5: Light wavelength spectrum emitted by two generic white LEDs. On the left the spectra of a white LED obtained by phosphors coating of a near ultraviolet LED, while on the right the spectrum of a white LED obtained by packaging together three different LEDs. Figure adapted from <http://zeiss-campus.magnet.fsu.edu/print/lightsources/leds-print.html>

The light spectrum may be an important factor to take into consideration when comparing a stimulus provided by different LEDs or by different devices, since col-

⁴The most common wavelength-converter materials are termed phosphors, which are materials that exhibit luminescence when they absorb energy from another radiation source.

ors perceived as similar may be given by very different light wavelength spectra ⁵. As will be detailed in Sec. 4.2.3, the SSVEP response is color dependent, but it is not straightforwardly related to the perceived color; previous studies in the vision research, using monochromatic lights, highlighted indeed a strong color/flickering-frequency inter-dependence.

COMPUTER DISPLAYS

To provide SSVEP eliciting stimuli, another commonly used device is a regular computer monitor. The major advantages of this kind of device are its wide diffusion, but also the fact that lot of the stimulus characteristics can be easily controlled by software.

Actually, compared to LED stimulator devices, it seems to be much more flexible, but it has also a major disadvantage; every computer monitor is able to update the image displayed on its screen at a certain frequency that is commonly set between 60 Hz and 85 Hz (the screen refresh frequency). This means that an upper limit to the displayable flickering frequency exist, given by the half of the screen refresh frequency for *flash* stimuli and by the screen refresh frequency for *pattern reversal* stimuli. Anyhow, much more importantly, since at every refresh only a single frame can be displayed for one whole period, there is a much stronger limitation to the displayable stimuli given by the fact that a device with framerate R can correctly render only frequencies of R/k , where $k \in \mathbb{N}$ for *pattern reversal* stimuli and where $k \in \mathbb{N} \geq 2$ for *flash* stimuli.

In the lower part of Fig. 4.1.6 is reported a schematic example of the stimulus resulting from the attempt to present a 24 Hz simple flash flicker on a regular 60 Hz display (where $60/24 \notin \mathbb{N}$), in comparison to the upper part of the figure where is reported the ideal stimulus that should be presented. This attempt results in an imprecise flickering stimulus that demonstrated to elicit weaker SSVEP responses, or at least SSVEP responses that are harder to detect [190].

There is not too much research works about the performance loss for SSVEP

⁵This phenomenon is known as metamerism and is due to the fact that the human eye contains only three different kind of photosensitive cells with different sensitivity for three different wavelength bands.

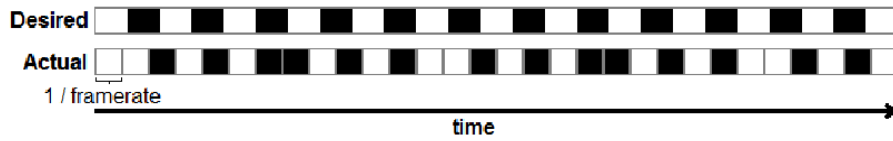


Figure 4.1.6: The difference between which states (black and white) are desired at each point in time and which states can actually be rendered. The length of each rendering alternates between being too long or too short, but is never quite correct. The example shown was derived from a 60 Hz device trying to render 24 Hz stimulation over the course of a half second. Figure taken from [10].

based BCIs given by an improper stimuli presentation as the one shown in Fig. 4.1.6; moreover in a lot of SSVEP based BCI implementations using computer monitors as stimulator devices, this issue is often ignored or underestimated. Despite of this, in various other works, researchers decided to use only frequencies that could be properly displayed. Furthermore in [190] has been demonstrated a strongly significant increase in the performances using properly displayable stimuli.

The same conclusion has been highlighted in a more controlled experiment also in [10], where the same LED stimulator device has been used to present an accurate square wave modulated stimuli (20 MHz accuracy) and a simulation of a 60 Hz monitor render of the same stimuli. In the performed experiment a strong performance increase has been shown when using properly rendered stimuli and moreover, when not doing so, has been highlighted the insurgence of various peaks in the EEG frequency spectrum which do not correspond to the fundamental stimulation frequency or its harmonics. These peaks means that the stimulus is eliciting responses synchronized on different frequencies than the desired one, lowering the needed frequency peaks, but moreover increasing the risk to interfere with the peaks elicited by other stimulation frequencies in the context of a multi-target *frequency tagged* SSVEP based BCI.

Therefore only correctly displayable stimuli should be used for optimal performances, although this impose severe limitations. For example in a regular computer monitor working at 60 Hz refresh rate there are only 6 correctly displayable frequencies greater than 8 Hz and moreover some of these frequencies are each

other harmonics ⁶, which is commonly undesirable for BCI applications, as will be later detailed.

Anyhow, not all the computer displays present the same limitations, concerning the screen refresh frequencies, high frequency displays exist and consumer devices can be found up to 120 Hz (e.g. displays implementing the NVIDIA 3D Vision Technology).

Being computer monitors much more flexible than LED stimulator devices in terms of presentable stimuli and moreover being much more handy in the implementation of human-computer interaction systems, various research works are available aimed at overcoming their limitations in terms of displayable targets. For example in [26] is presented a technique to discriminate the targets not only according to their flickering frequency, but also to their duty cycle, as will be explained in Sec. 4.2.2. In [194] is proposed to use a stimulus using a varying duty cycle in order to maintain a stable frequency presentation and avoid the instabilities showed in Fig. 4.1.6. Various works are available also regarding the discrimination of the phase of the stimulus instead of its frequency, as will be detailed in Sec. 4.3.3. Moreover, recently have been proposed also the use of mixed frequencies in the same target [76].

Apart from the screen refresh frequency, other displays features should also be considered for SSVEP stimuli presentation. Three main technologies of consumer computer monitors are available on the market: Cathode Ray Tube (CRT), Liquid Crystal Displays (LCD) with Cold Cathode Fluorescent Lamp (CCFL) back-light and recently also LCD with LED back-light. Indeed a study has been performed in order to compare their performances for SSVEP stimuli presentation [198].

CRT computer monitors are based on the Cathode Ray Tube technology that is mainly a vacuum tube with three electron guns (one for each color) at its rear, shooting electrons on a flat screen where fluorescent materials are positioned. Three materials are commonly used in order to emit red, green and blue light when hit by electrons. A problem which may arise using CRT displays, in the context of SSVEP stimuli presentation, is due to the fact that each electrons gun is able to shot only one beam of electrons at a time, thus to hit only one screen location (or pixel). At every refresh, to display a new frame, the electrons beam has to be moved over

⁶In particular the frequencies are: 8.57, 10, 12, 15, 20, 30.

every pixel of the screen, one at a time, scanning repetitively and systematically in a fixed pattern called a raster⁷ the entire front area of the tube. The electrons gun in this scanning operation has therefore to “paint” the whole frame at a particular scanning frequency, which is commonly fast enough to be not perceived, but can be clearly seen when a beating between different frequencies is present (e.g. looking on a television the video recording of another television). Indeed as demonstrated in [198], although CRT monitor can be used, the elicited SSVEP response presents various components other than the fundamental components and its harmonics, which are probably due to the screen refresh mechanism and the frequency beatings with the stimulation flicker.

CCFL-LCD and LED-LCD are based on a completely different technology, where commonly a liquid crystal panel is positioned in front of a lamp. The panel is able to modulate the strength of three colored light filter at each pixel location, thus modulating the amount of red green and blue light able to reach the user eye for each pixel. LCDs lit for each frame all the pixels at once, consequently they do not present the same problem as CRT displays. On the other side, it is harder to implement high frequency LCDs and most of them have a fixed refresh frequency of 60 Hz. The main difference between CCFL and LED back-light is the used lamp technology, thus it may considerably change the light wavelength spectrum of the three primaries. Considerable differences in the light wavelength spectrum are also present with respect to CRT displays, whose spectrum is reported in Fig. 4.1.7, which in turn is considerably different from the LED ones presented in Fig. 4.1.5.

The spectral characteristic of the RGB primaries of a computer display (or of an RGB LED) define a color space of possible representable colors that is commonly referred as a color gamut [21]. The representable colors by a computer display are commonly restricted to a subset of all the colors perceptible by the HVS. Thus, as will be detailed in Sec. 4.2.3, also the color gamut may influence the effectiveness of the presented stimulus in terms of the elicited SSVEP response strength.

⁷The word “raster” has been later borrowed in the world of computer graphics for this reason.

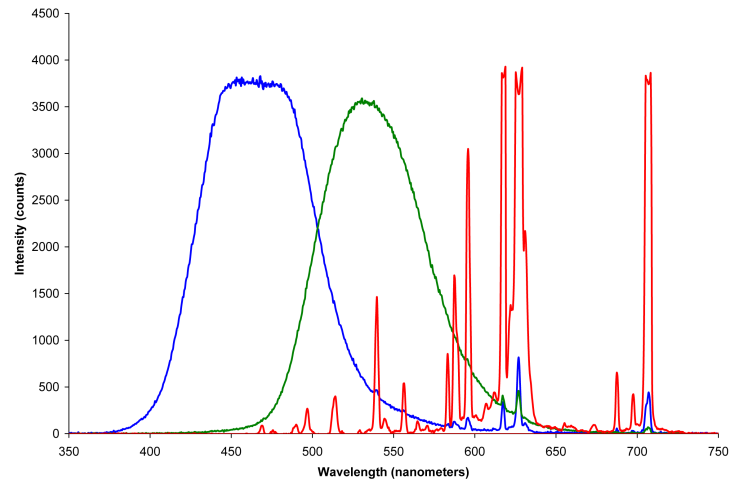


Figure 4.1.7: Spectra of individual color phosphors of a typical CRT video monitor. Note that there is some overlap of very strong red phosphor spectral peaks onto the other spectra probably due to electrons "bleeding over" into adjacent colored phosphor dots, since focusing of the electron beams is not perfect in a CRT. Figure courtesy of <http://en.wikipedia.org/>.

4.2 RESPONSE CHARACTERIZATION

In general, as previously mentioned, a SSVEP response may be generated by any Repetitive Visual Stimulus (RVS) changing one of its properties, e.g. color and pattern, but also shape, position, stereoscopic depth, etc. at a specific frequency.

The amplitude of the SSVEP response, measured by an EEG device, changes according to the used stimulation frequency and in general it decreases as the frequency increase, with three local maxima.

From previous studies [157, 188] it is known that these three local maxima are related to different neural subsystems which are more sensitive to a particular stimulation frequency band. The three main components can be observed in average across different subjects, as shown in Fig. 4.2.1, although a considerable subject dependance and stimulus dependence have been highlighted.

It is common to distinguish the three different components associated to the three peaks in Fig. 4.2.1 as *low*, *medium* and *high* frequency components. Different components have different relative amplitudes, but also different latencies, given

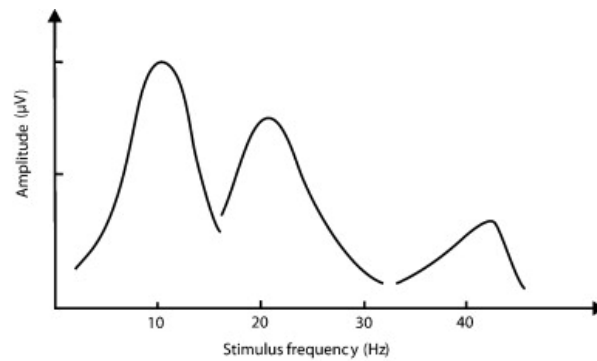


Figure 4.2.1: Amplitude of the SSVEP response with respect to the stimulation frequency, highlighting the three main components. Synthetic model. Figure adapted from [188].

by the time needed to reach a frequency-stable response from the first stimulus on-set; according to [188]:

- The high-frequency component (peak on the right in Fig. 4.2.1), in the gamma range, is characterized by a small interindividual variability and latencies of about 30 ms to 60 ms.
- The medium-frequency component (peak in the middle of Fig. 4.2.1) in the 15 Hz to 25 Hz range with higher interindividual variability and latency of about 85 ms to 120 ms.
- The low-frequency component (peak on the left in Fig. 4.2.1), below 15 Hz with the higher latency of 135 ms to 350 ms

According to [157], but also to more recent studies [188], responses generated in these three ranges seems to involve different cortical areas. Anyhow, although this source is not the only responsible for the SSVEP generation, most of the experimental data proved that for all the three components the strongest local source of SSVEPs is located in the primary visual cortex (also known as striate cortex V1, equivalent to Brodmann area 17) in the occipital region of the brain.

The considerable subject/stimulus dependance of the SSVEPs amplitudes in the three different frequency bands can be appreciated in Fig. 4.2.2 where are depicted the results of an experimental evaluation, presented in [193], of the three

main components which shows quite different peaks than the ones presented in [188] and showed in Fig. 4.2.1. Also according to the experiments performed in [143] the amplitude of the SSVEP response for most of the subjects seems to have a global maximum around a stimulation frequency of 15 Hz, counterintuitively with respect to Fig. 4.2.1.

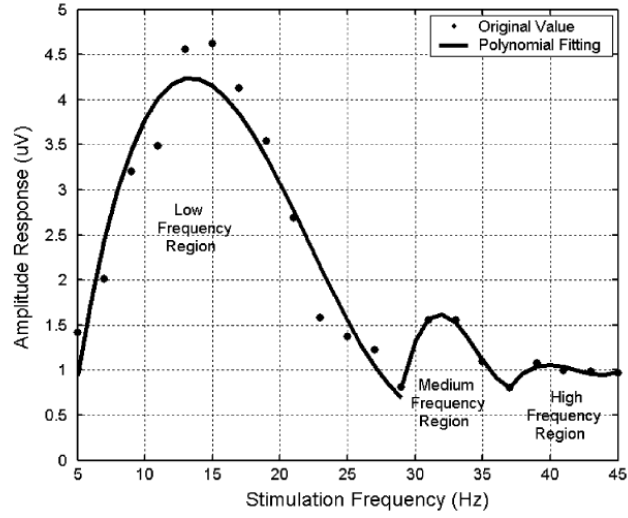


Figure 4.2.2: Amplitude of the SSVEP response with respect to the stimulation frequency, highlighting the three main components. Points are real data acquired from a real subject, while lines are polynomial interpolations. Figure adapted from [193].

Indeed, also as stated in [188], despite years of investigation, the complex mechanisms behind SSVEPs are not yet fully understood. As stated in [193], but also known since [157], the amplitude of the SSVEPs varies in a complex manner with the frequency of stimulation, according to the luminance of the stimuli, its spatial frequency, its flickering modulation depth and its color (i.e. wavelength spectral distribution).

Therefore in this section will be given a review of previous works aimed at shading light on the SSVEP characterization and thus to provide essential information to chose the stimulation frequencies and stimuli properties for efficient SSVEP based BCIs.

HARMONICS COMPONENTS

The SSVEP response does not appear as a simple sinusoid in the EEG signal, but is given by the quasi-sinusoidal response at the fundamental frequency corresponding to the stimulation frequency, summed with multiple quasi-sinusoidal signals corresponding to harmonics and in some cases subharmonics. In Fig. 4.2.3 is sketched a representation of the observed harmonic and sub-harmonic components given a single stimulation frequency.

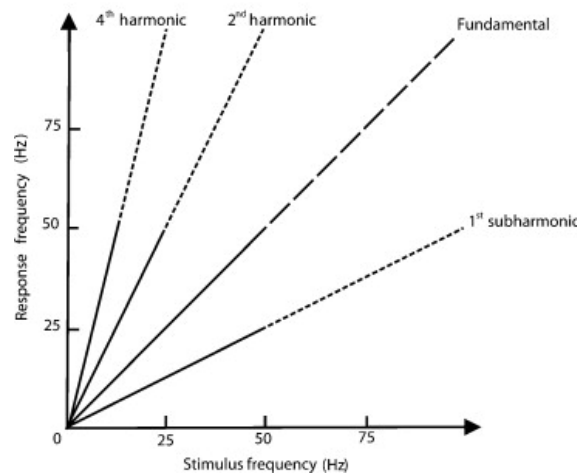


Figure 4.2.3: Harmonics of the SSVEP response that have been observed with respect to the stimulation frequency. Figure adapted from [188].

The intensity of each component is not reported in Fig. 4.2.3, but is commonly higher for the fundamental frequency and descending for the subsequent harmonics, while the sub-harmonic is commonly much weaker.

For SSVEP based BCIs, considering also the harmonics frequency response as features for the classification, is known to increase the detection speed and also the accuracy, thus provides an higher ITR [53, 92].

A study has also been performed to identify how many harmonic components is worth taking into account for BCIs applications with respect to the used stimulation frequency [55], showing that better results can always be obtained using at least the first harmonic and for some frequencies it is worth to use also the second and third one.

Anyhow, as will be introduced in the next sections, also the presence and amplitude of higher/lower harmonics of the fundamental frequency depends on the provided stimulus; thus the results presented in [55] are probably valid only for the same (or a similar) kind of stimulus.

4.2.1 STIMULUS SIGNAL SHAPE DEPENDANCE

The presence of different harmonics of the stimulation frequency is not a well understood phenomenon and is considered an evidence of non-linear dynamics as will be introduced in Sec. 4.2.8.

Despite of this, the intensity of the harmonic components in the SSVEP response has been recently linked to the stimulation signal shape by a study that has been conducted in order to compare different stimulation signals to determine which one elicits the stronger SSVEP response for BCIs applications [181]. In this context, for “stimulation signal shape” is meant whether the flickering stimulus is modulated by a square-wave signal (i.e. it is lit on and off repetitively), a sinusoidal signal (i.e. its intensity is smoothly increased and decreased following a sinusoid), a sawtooth signal, etc.

This study [181] demonstrated that the presence and intensity of the harmonic components is correlated to the stimulation signal shape and that a square wave modulated flickering stimulus is the best choice if a stronger response is the objective. This was already noticed and mentioned in [193], although no experimental data showing different stimuli was reported.

On the other side, a sinusoidal modulation of the stimulus should be preferred if the objective is to have fewer or weaker harmonics. This may be useful if higher harmonics of a lower frequency could interfere with higher frequencies used to implement multiple targets *frequency tagged* SSVEP based BCIs.

4.2.2 DUTY CYCLE DEPENDANCE

In the aforementioned research work [181], but also previously in [26] and more recently also in [74], the effect of the duty cycle of the flickering stimulus has been analyzed in order to estimate its influence on the SSVEP response.

Despite of the fact that in most of the works the information about the stimu-

lus duty cycle is not taken into consideration, it demonstrated to influence (with different intensities for different stimulation frequencies), the amplitude of the SSVEP response at both the fundamental frequency and its harmonic components. Consequently it has been demonstrated also that it influences the SSVEP detection accuracy for SSVEP based BCIs applications, giving better accuracies with duty cycles between 0.4 and 0.8 (as mentioned also in [193]), for most of the tested frequencies, in various research works.

More interestingly, the effect of the duty cycle on the harmonic components of the SSVEP response has been successfully exploited also to discriminate between different targets flickering at the same frequency, for SSVEP based BCIs applications [26]. This has been possible thanks to the fact that different duty cycles modulates differently the SSVEP fundamental and harmonics responses, leading to the possibility of classifying them according to their relative amplitudes.

This is an interesting result in view of the limited frequencies available on regular computer displays to present flickering stimuli, as previously addressed in Sec. 4.1.3.

Another interesting research work [194], proposed instead the use of a duty cycle modulation aiming to overcome the aforementioned frequency limitation imposed by computer displays. The main idea is to renounce to the duty cycle stability in exchange for stable flicker frequencies which otherwise could not be properly displayed due to the monitor refresh frequency, as shown in Fig. 4.2.4.

4.2.3 COLOR DEPENDANCE

The mechanisms underling the human color vision have been deeply studied in the past in the field of vision research [205], but also regarding this topic researches are still ongoing.

Perception of color begins with specialized retinal cells containing pigments with different spectral sensitivities, known as *cone* cells as introduced in Sec. 2.1.3. In humans, there are three types of *cones* sensitive to three different spectra, resulting in trichromatic color vision.

The *cones* are conventionally labeled according to the ordering of the wavelengths of the peaks of their spectral sensitivities: short (S), medium (M), and long (L) *cone* types. These three types do not correspond well to particular colors as we

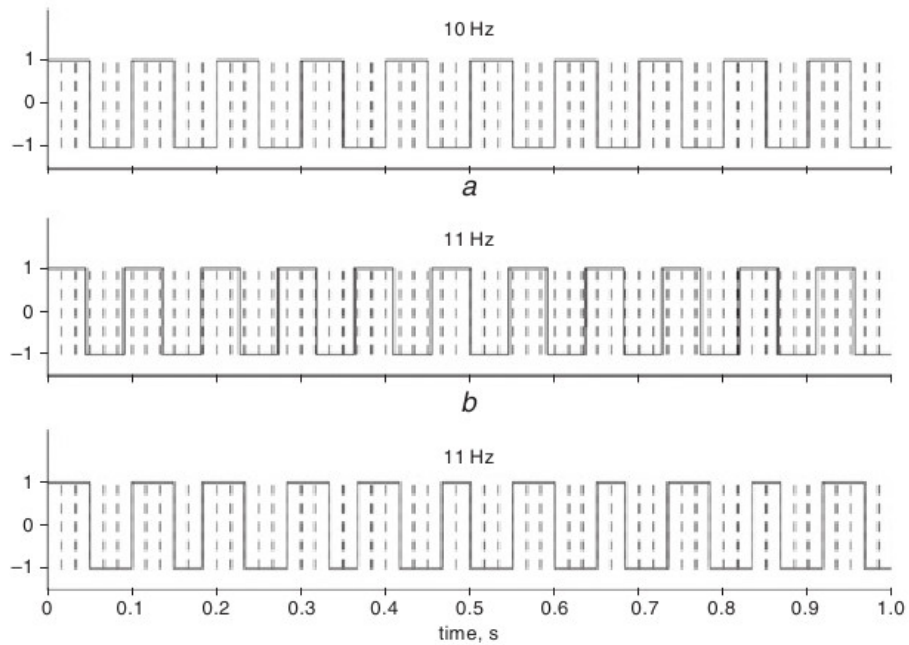


Figure 4.2.4: Dashed line represents the 60 Hz framerate of a regular computer monitor. In (a) is represented a 10 Hz square wave signal modulating a flickering stimulus that could be properly displayed on a 60 Hz computer monitor. In (b) is represented a 11 Hz square wave signal modulating a flickering stimulus that could not be properly displayed on a 60 Hz computer monitor. In (c) is represented a 11 Hz square wave signal modulating a flickering stimulus with a variable duty cycle in order to maintain a stable 11 Hz also on a 60 Hz computer monitor. Figure taken from [194].

know them. Rather, the perception of color is achieved by a complex process that starts with the differential output of these cells in the retina and it will be finalized in the visual cortex and other associative areas of the brain.

Anyhow, in simple words, we may say that *S-cones* are more sensitive to bluish light, *M-cones* to greenish light and *L-cones* to reddish light, as shown in Fig. 4.2.5. Thus, three parameters, corresponding to levels of stimulus of the three types of cone cells, can in principle describe any color sensation.

The light that commonly enter our eyes is not monochromatic light (e.g. light of a single wavelength), but light composed by a specific spectra that having more or stronger components in one or more of the S, M, L bands is perceived as being of different colors. The same perceived color may be due to light with very different

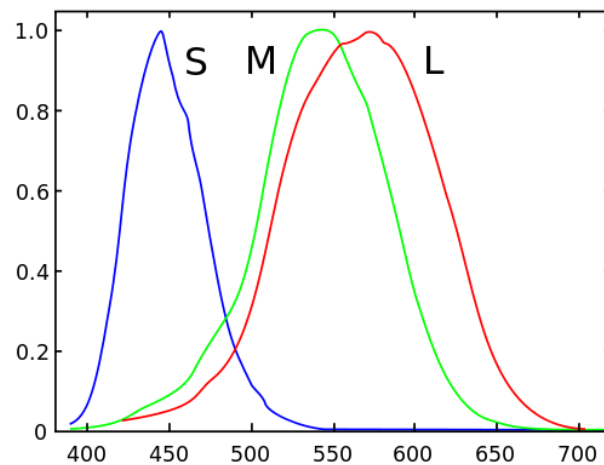


Figure 4.2.5: The normalized spectral sensitivity of human cone cells of short-, middle- and long-wavelength types. Figure taken from <http://en.wikipedia.org>.

spectra⁸, since for color perception what is important is not the spectrum, but the spectral power distribution weighted by the S, M, L spectral sensitivity curves⁹.

Interestingly, in a study conducted at the beginning of the research on SSVEP responses [155], has been discovered that the SSVEP response amplitude depends on the stimulation light spectrum. In particular, in the performed experiment has been used a stimulation device able to present simple *flash stimuli* of an almost monochromatic light (being comparable with nowadays LED light stimulator devices). The used stimulator could present three different lights characterized by a different peak in their wavelength spectrum, respectively centered in 435 nm, 589 nm, 634 nm. Changing the flickering frequencies, three different plots of the peak-to-peak SSVEP response amplitudes have been produced, showing significantly different curves, as reported in Fig. 4.2.6.

This result demonstrates that the SSVEP responses have different amplitude-frequency characteristics for different light spectra.

This is an extremely important fact that may explain a lot of the contrasting

⁸This phenomena is named *Metamerism*. Defined as the matching of apparent color of objects with different spectral power distributions. Colors that match this way are called metamers.

⁹Spectral sensitivity curves for an average observer have been experimentally computed and are commonly referred as *CIE Color matching functions*

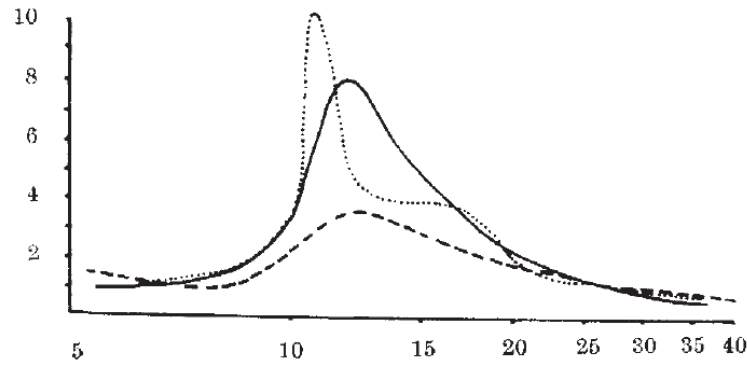


Figure 4.2.6: Peak-to-peak amplitude (μV) of the SSVEP response versus the stimulus modulation frequency (Hz) for a subject. Solid line for the 435 nm (blue) stimulation light, dashed line for the 589 nm (yellow) stimulation light and dotted line for the 634 nm (red) stimulation light. Figure adapted from [155].

results found in various research works investigating the performance of different stimuli for SSVEP based BCIs. It may explain also the different amplitude-frequency characteristic presented in Sec. 4.2.

Moreover, as detailed in Sec. 4.1.3, a green light LED stimulator would provide a completely different spectrum than a CRT monitor displaying a perceptively equal green flash stimuli, due to metamerism; anyhow this possible influence seems never to be taken into account in recent SSVEP based BCI implementations.

A more recent study, aimed to estimate the impact of the stimulus color on the performance of a SSVEP based BCI, discovered that on an LCD monitor a significant performance increase can be obtained using a white stimulus than a colored one [24]. This could be simply explained by the luminance (and thus contrast) difference which seems not to have been equalized in the cited work. Otherwise, according to the results reported in [155], it may be due to the fact that white light has a spectra with a power distribution spanning over a wider wavelength band and thus probably able to hide the specific SSVEP response characteristics provided by monochromatic lights (as the ones shown in Fig. 4.2.6).

Similar results have been showed in [10, 11] reporting as the best stimuli, in terms of elicited response strength, a white patch over a black background.

Although the SSVEP response amplitude to a stimulus having a spectra able to

stimulate all of the three kind of *cones*, would hardly be a simple linear sum of the effects of the different singular monochromatic components, it is probable that an almost flat light spectrum (i.e. with an high number of components) would provide a flatter amplitude-frequency characteristic given by a weighted sum of the different components.

Anyhow, the selection of the stimulus color has to be designed taking into account its influence on the performances, looking for a compromise also for the users' comfort [34]. Moreover, in my opinion, also the light spectrum provided by various devices should be taken into consideration.

Further experiments should be performed in this direction to analyze the significance of the impact of different light spectra on the performance of SSVEP based BCIs.

4.2.4 STIMULUS SIZE DEPENDANCE

It is known from previous studies in the field of vision research that also the flickering stimulus size is modulating the SSVEP response intensity [23]. In particular it is known that stimulating a bigger area of the user field of view, a stronger response could be observed, probably due to the activation of an higher amount of retinal cells (*cones* and *rods*) and thus also of bigger neural populations along the visual pathways.

Recently more studies have been conducted in order to estimate how much this modulation due to the stimulus size could affect the detection accuracy in view of SSVEP based BCI applications [133]. According to these results a significant detection accuracy improvement can be obtained using flickering stimuli subtending at least 2° of the user visual angle, while smaller stimuli elicit weaker responses. Stimuli subtending more than 2° of visual angle seems to still elicit slightly stronger response as the stimuli became bigger, but the increase in accuracy is not much significant [133].

Stimulus size is known to change also the SSVEP response with respect to the stimulation color [161], since *cones* are mainly concentrated in the foveal area of the retina, while *rods* are not. Consequently different stimulus sizes stimulate in different proportions *rods* and *cone* cells.

4.2.5 INTER-STIMULI DISTANCE

In the same research work [23], also the inter-stimuli distance is taken into consideration with respect to the detection accuracy in view of BCI applications.

It is known from past works in vision research that the evoked potential response is greatest for light stimuli entering the fovea and decrease as a Gaussian function of width 5° centered at the point of fixation [35]. Thus in [23] have been evaluated the impact of the distance between the different flickering stimuli on the SSVEP response detection accuracy.

Indeed as expected a target separation of at least 5° of visual angle or more is needed for optimal SSVEP based BCIs.

4.2.6 STIMULI NUMBER

More interestingly, in the same work [23], also the number of different flickering targets showed to influence the SSVEP response detection accuracy in conjunction to the target distance. At a target distance of 2° the increase of the number of targets (where every target flicker at a different frequency), decreases the SSVEP response detection accuracy, as expected, since multiple flickering frequencies are entering the foveal area. On the other side, if the target distance is kept higher than 5° , an increase of the number of targets increases the SSVEP response detection accuracy.

Although may at first seems counterintuitive, this finding is supported by previous researches suggesting that visual processing is a limited resource, hence the interference caused by competing stimuli is reduced as perceptual load (e.g. the number of competing flickering stimuli) is increased [97].

From the BCI research point of view this is a very interesting result, since having an high number of targets is important to obtain an optimal ITR, as to have an high detection accuracy.

4.2.7 SPATIAL FREQUENCY DEPENDANCE

When using checkerboard patterned stimulus (or any other patterned one) also the spatial frequency of the stimulus is involved in the VEP and SSVEP response.

In the case of a checkerboard *patterned stimulus*, an high spatial frequency translates to small checks, while a low spatial frequency translates to bigger checks. Consequently, simple *flash stimulus* could be seen as a *patterned stimulus* with zero spatial frequency, while a zero-contrast stimulus could be seen as a *patterned stimulus* with infinite spatial frequency.

As mentioned in Sec. 4.1.1, in [138], for *patterned stimuli* is recommended to report the check size in terms of the subtended visual angle with respect to the user; thus also for the spatial frequency, spatial changes have to be reported in terms of changes per degree of visual angle.

Various experiments have been conducted to assess an optimal spatial frequency to elicit stronger SSVEP responses, but an high subject variability has often been reported. In [10, 11] it is shown that the spatial frequency of the stimuli influence the SSVEP response in a non-linear fashion, but locally optimal responses could be achieved with purely *flash stimuli* (thus a a spatial frequency of zero) and with a *patterned stimulus* with a spatial frequency of 6.5 alternations per degree.

4.2.8 NON LINEARITY

Although the SSVEP response, being a frequency-locked and phase-locked response to a flickering stimulus, may seems a linear mechanical activation of large group of neurons in response to light stimuli, it is known not to be linear since long time [157] and various evidences suggest it to be governed by far more complex dynamics.

Indeed, various experiments in different research fields confirmed the presence in the SSVEP response recorded by EEG electrodes of signal components elicited by the stimuli presentation that could not be possible assuming the HVS as a linear system, from a signal processing point of view. In [33] the HVS non-linearity is reported to have at least seven orders, although it is not too clear how the value has been inferred.

The most used tools to study non-linear behaviors in the SSVEP response have been Higher Order Spectra (HOS) analysis tools and in particular the Bispectrum [75] which in contrast to second-order spectral analysis take into account the interactions between each harmonics.

Also recently, in BCI research, non-linear harmonics coupling are frequently reported and defined as unclear [181], but interestingly it seems not yet to exist any SSVEP based BCI using HOS analysis, although it has been reported to achieve good results in the EEG signals analysis for clinical applications [75] and is also commonly used for anesthesia depth monitoring (e.g. the BIS index). Further studies in this direction for BCI applications have been encouraged also in [188].

4.2.9 AFFECTIVE MODULATION

Another evidence suggesting the SSVEP response not being only a mechanical reaction of the brain to a flickering stimulus, is that it is modulated by the stimulus semantic and affective content, both in its amplitude, latency and topographical propagation. Results in this direction have been obtained using as stimuli flickering pictures instead of solid color patches or checkerboards, showing a clear correlation between the stimulus semantic and the evoked SSVEP response [87, 183].

Since it is known that the SSVEP response is heavily dependent on the attention the subject pays to the flickering stimuli, these results could be due to an increase of the interest elicited in the user by the emotional content of the stimuli.

Exploiting these observations, the implementation of Affective BCIs (aBCIs) based on the SSVEP modality has been foreseen in [57, 119] on the wake of aforementioned results in the field of neuroscience [83, 85] and of some recently proposed proofs of concept [6, 195], which demonstrated the enhancement of the SSVEP response power using pleasant or unpleasant emotional flickering pictures as happy or angry human faces.

An aBCI based on the SSVEP modality could be used to assess the subject attention or affective state with respect to different flickering stimuli as pictures, objects in a VR environment or entities in a Computer Game.

Despite of this, as far as I know, SSVEP based BCIs were never used in the context of aBCI; probably mainly because of the fact that in contrast to other modalities, flickering stimuli have to be provided and that the flickering itself may fictitiously divert the user attention. Moreover from the commonly used signals analysis techniques for SSVEP based BCIs, only one degree of freedom could be extracted: the SSVEP response power.

On the other hand, for some applications, the mentioned limitations could be overcome. Although the presentation of stimuli is a serious limitation for the implementation of aBCIs in several context, it is not in the field of VR environments, Computer Games or Augmented Reality (AR), where the subject is commonly looking at a synthetic, or partially synthetic world, where flickering stimuli could be included as detailed in Sec. 4.1.3 and implemented in Sec. 5.4.2.

Moreover, the user attention diversion caused by the flickering stimulus could be minimized, using higher frequencies that are less consciously perceptible, or it could be exploited, as will be proposed in Sec. 6.3. Another option would be to employ various flickering stimuli at the same frequency, but with different phases in order to obtain the same “attention diversion” due to the flickering, but to be able to detect which target is modulating the strongest SSVEP response. The detection of phase modulated stimuli will be addressed in Sec. 4.3.3.

Furthermore, using more complex signal processing techniques taking into account the SSVEP response propagation from the occipital to the parietal and frontal areas of the cerebral cortex, as proposed in [119], more information about the valence and/or arousal of the emotion involved in the SSVEP response elicitation could be deduced. It is in fact demonstrated that the modulation of the SSVEP response, due to the user’s affect state, changes across different scalp locations in correlation with the arousal and valence of the elicited emotion [85].

4.3 SIGNAL ANALYSIS

The EEG signals processing to detect the SSVEP response could be very simple. Lot of research works, published so far, indeed simply apply to raw EEG signals’ epochs a Fast Fourier Transform, in order to evaluate them in the frequency domain and to estimate the signal power in the frequency regions corresponding to the flickering frequency of the stimuli.

Multiple commands, for BCI applications, are commonly associated to targets flickering at different frequencies, thus the variation in signal power in the different frequency bands gives the information on the attended target. This is commonly referred to *frequency tagging* to differentiate it with respect to another technique named *phase tagging* which will be addressed in Sec. 4.3.3.

Scope of the signal analysis is therefore to compute a feature describing the intensity of the SSVEP response to a stimulus flickering at a particular frequency, in order to determine the attended target.

In terms of digital signal processing, a multichannel EEG signal epoch (i.e. a slice of signal acquired in a time window $[t, t + \Delta t]$) can be represented as a matrix \mathbf{X} of size $T \times N$ where the N columns corresponds to the channels (i.e. signals coming from the different electrodes) and T are the samples for each channel.

The most trivial approach is to compute the Fast Fourier Transform (FFT) on each of the channel signal \mathbf{x}_n epochs and then evaluate the frequency bins height for each of the stimulation frequencies.

A threshold could be set for each of the stimulation frequencies in order to decide if a SSVEP response is present or not, since in different frequency bands the responses may have very different amplitudes. A calibration phase to set the best thresholds is a common approach and is equivalent to train a linear classifier using the power in the different frequencies as features. The same operation could be applied also to a chosen number of harmonics for each of the stimulation frequencies.

One of the limitations of this approach is given by the fact that to achieve an higher frequency resolution in the frequency domain, a longer signal epoch in the time domain has to be used, since the frequency resolution is given by $1/\Delta t$.

Moreover the computed power in each frequency/harmonics do not contains only the SSVEP response, but also all the underlying stimulus-uncorrelated brain activity occurring at the same frequency. Thus what is obtained is an absolute value of the neural activity in a specific frequency.

These may not be major issues when using long signal epochs and/or multiple trials, as was commonly done in clinical applications or in vision research. Nevertheless, for BCI applications, to achieve high ITRs, epochs as short as possible have to be used and multiple trial approaches have to be avoided.

Another approach which is similar, but do not use the FFT algorithm, is to estimate the whole signal energy after applying a narrow-band filter. The SSVEP response to a stimulus flickering at a given frequency f could be estimated applying a narrow band FIR filter centered around f to the EEG signal $x(t)$.

The energy of a generic signal $x(t)$ in the time domain can be computed as:

$$E = \frac{1}{\Delta t} \int_t^{t+\Delta t} |x(t)|^2 dt \quad (4.1)$$

Being \mathbf{x}_f a discrete time signal, obtained applying to the samples \mathbf{x} acquired by the EEG device a narrow band filter centered around f , the signal energy can be computed as:

$$E_f = \frac{1}{T} \sum |\mathbf{x}_f|^2 \quad (4.2)$$

This method is computationally faster than the FFT method, but again E_f do not contain only the SSVEP response at the stimulation frequency f , but also all the underlying stimulus-uncorrelated brain activity occurring at that frequency.

Despite of the fact that the methods mentioned so far have been successfully used both for clinical applications and for SSVEP based BCI implementations, they present several issues which limit the obtainable detection speed and accuracy.

In the context of SSVEP based BCIs, the limitations imposed to the number of usable stimuli frequencies by ordinary displays in conjunction to the usage of stimuli privileging users' comfort, cause weaker SSVEP responses which are harder to detect. Moreover, in contrast to most of the clinical applications, for BCI applications the timing constraint is strongly relevant, thus a BCI command has to be detected in less than 4 s to 5 s to be usable and in less than 2 s to 3 s to be perceived as a real-time control.

Therefore, the need to detect weaker SSVEP responses in real-time for BCIs applications, in the last years pushed the researchers to move on more complex signal processing techniques trying to exploit all the information contained in the SSVEP response and trying to overcome some of the limitations imposed by some stimulator devices.

In this Section the state-of-the-art techniques developed in the last years will be reviewed focusing mainly on the methods that revealed to provide the best results. Moreover, will be described the methods used to mix the signals coming from the different electrodes in order to extract most of the available information regarding the SSVEP response. Furthermore, different methods to separate the SSVEP

waveform from the uncorrelated brain activity happening in the same frequency will be introduced.

At the end will be presented also the *phase tagging* technique to implement SSVEP based BCIs using targets flickering all at the same frequency.

4.3.1 SPATIAL FILTERS

In EEG signal recording is common to utilize multiple electrodes and in various clinical applications (or some BCI modalities as well) this is needed to obtain information about the spatial origins of the recorded signal features. Despite of this, in some applications, in particular in the context of BCIs, information coming from multiple electrodes may need to be “summed” in order to estimate a single feature (e.g. the SSVEP response intensity) from a set of electrodes or from the whole electrodes set. This is the case in particular for the SSVEP response detection [53, 55].

In this case, the best performing method to sum all the electrodes contribution in order to estimate the signals features has to be found. The linear combination of the signals recorded at different scalp sites is usually referred as *spatial filtering* and several methods have been investigated in the field of BCI research, where a fast and reliable detection is crucial for interactive applications.

Trivial solutions as summing all the channels’ signals together in the temporal domain could give very bad results, since the SSVEP response may be significantly phase shifted across different brain regions [30, 55]. Consequently an average between different channels’ signals may cancel out the SSVEP waveform from the resulting sum, as is depicted in Fig. 4.3.1.

Another trivial approach, which otherwise have been successfully used, is to choose on a per-subject basis the best bipolar combination of channels. This is accomplished looking for the SSVEP response on all the signals obtained as differences between couples of channels’ signals, trying all the possible combinations to choose the one where the stronger response could be read. Finding a bipolar combination between two channels’ signals, means to look for two locations where the SSVEP responses are in counter-phase and thus where a subtraction between the signals would enhance the SSVEP waveform in the resulting signal. This approach

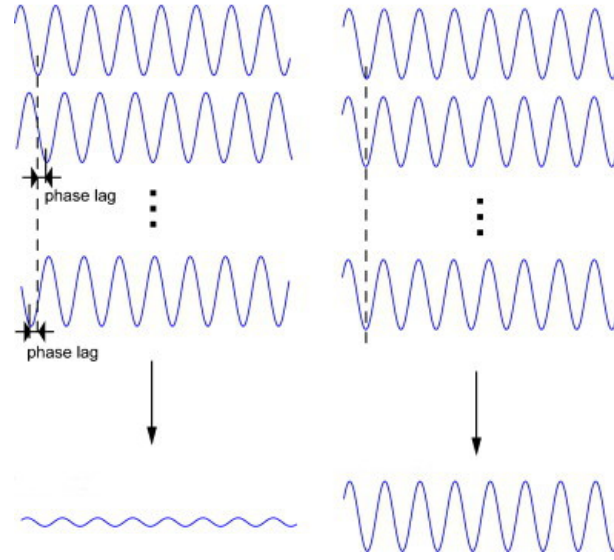


Figure 4.3.1: Synthetic figure showing two ideal SSVEP waveform hypothetically acquired by different scalp locations (in the upper part of the figure). On the left the two waveforms are not in phase, while on the right they are. Starting from equally intense responses, in the bottom are shown the different results obtainable with a trivial spatial filtering (time-domain averaging) in the two cases. Figure adapted from [188].

has also the advantage of reducing significantly the artifacts since the two chosen electrodes are commonly very close to each other. One of the bipolar combinations giving the best results for most of the subjects in different research works has been reported as the Oz-Cz combination [10, 207].

Another approach is to compute the SSVEP response power, or any other kind of feature for each channel, to later add the results together using different methods according to the kind of features.

Anyhow the procedure which gave the best results is to compute a spatial filter in order to merge all the channels' signals in the temporal domain avoiding to elide the SSVEP waveform, but otherwise trying to assign different weights to the contribution of each channel's signal in order to enhance the SSVEP waveform in the result [55].

The signal obtained by applying a spatial filter \mathbf{w} to the multichannel signal \mathbf{X} can be written as $\mathbf{x}_w = \mathbf{X}\mathbf{w} = \sum_{n=1}^N w_n \mathbf{x}_n$ where w_n is the n -th column of \mathbf{X} (and

thus the n -th channel recording) and w_n the n -th spatial filter coefficient. The goal of a spatial filtering algorithm is to find a spatial filter \mathbf{w} maximizing the signal components designed as features with respect to the signal components considered as noise or background brain activity. Otherwise, the goal may also be to maximize the signal variance between the features that will be used to discriminate between different conditions.

Starting from the Motor Imagery paradigm [152], the use of spatial filters was later introduced also in the context of SSVEP based BCIs and demonstrated to improve the SSVEP detection, reducing the signal epochs length needed to reliably detect a SSVEP response and thus increasing the information transfer rate [55].

In particular for the SSVEP detection various methods have been proposed and here will be discussed two of them: the Common Spatial Patterns (CSP) method which is the default one used in the OpenVibe software ¹⁰ (described in Sec. 5.3) and the SSVEP Minimum Energy Combination method [53], which has been chosen in this work as one of the best performing methods [55].

COMMON SPATIAL PATTERNS

The main idea behind the CSP method is to compute a spatial filter able to linearly mix the signals coming from electrodes in order to maximize the variance between two conditions. In other words, the filter aims is to project the acquired signals in a lower dimensional space, where the maximum variance between the two conditions is reached in order to ease the work of a classifier of recognizing the two different classes.

The CSP was previously used for the detection of abnormal EEG patterns in clinical applications and has later been introduced in the context of BCIs for the Motor Imagery paradigm [152] and in particular for the discrimination of left-/right actual and imagined hand movements.

This method is based on the simultaneous diagonalization of two covariance matrices; recalling the notation presented at the beginning of Sec. 4.3.1, given a multichannel EEG single trial recording \mathbf{X} , its normalized spatial covariance can

¹⁰<http://openvibe.inria.fr/steady-state-visual-evoked-potentials/>

be obtained by:

$$\mathbf{C} = \frac{\mathbf{X}\mathbf{X}^\top}{\text{trace}(\mathbf{X}\mathbf{X}^\top)} \quad (4.3)$$

where $\text{trace}(\mathbf{x})$ is the sum of the diagonal elements of \mathbf{x} . For both the conditions (e.g. trials of imagined left hand movements and right hand movements), the covariance matrix $\overline{\mathbf{C}}_{g \in [A, B]}$ is computed by averaging over the trials of each group A and B , then the composed spatial covariance is given by $\mathbf{C}_c = \overline{\mathbf{C}}_A + \overline{\mathbf{C}}_B$ and it can be factored as $\mathbf{C}_c = \mathbf{U}_c \boldsymbol{\lambda}_c \mathbf{U}_c^\top$, where \mathbf{U}_c is the matrix of eigenvectors and $\boldsymbol{\lambda}_c$ is the diagonal matrix of eigenvalues. Later is applied the whitening transform to equalize the variance in the space \mathbf{U}_c letting all the eigenvalues of $\mathbf{P}\mathbf{C}_c\mathbf{P}^\top$ equal to one:

$$\mathbf{P} = \sqrt{\boldsymbol{\lambda}_c^{-1}} \mathbf{U}_c^\top \quad (4.4)$$

Moreover, if $\overline{\mathbf{C}}_A$ and $\overline{\mathbf{C}}_B$ are transformed as: $\mathbf{S}_A = \mathbf{P}\mathbf{C}_a\mathbf{P}^\top$ and $\mathbf{S}_B = \mathbf{P}\mathbf{C}_b\mathbf{P}^\top$ then \mathbf{S}_A and \mathbf{S}_B share the same common eigenvectors and if $\mathbf{S}_A = \mathbf{D}\boldsymbol{\lambda}_a\mathbf{D}$ then $\mathbf{S}_B = \mathbf{D}\boldsymbol{\lambda}_b\mathbf{D}$ and $\boldsymbol{\lambda}_a + \boldsymbol{\lambda}_b = \mathbf{I}$, where \mathbf{I} is the identity matrix.

Since the sum of two corresponding eigenvalues is always one, the eigenvector with largest eigenvalue for \mathbf{S}_A has the smallest eigenvalue for \mathbf{S}_B and vice versa. This property makes the eigenvectors useful for classification of the two distributions. The projection of whitened EEG onto the first and last eigenvectors will give feature vectors which are optimal for discriminating two populations of EEG epochs in the least squares sense [152].

With the projection matrix $\mathbf{W} = (\mathbf{D}^\top \mathbf{P})^\top$ the application of the filter to every recorded EEG trial is given as $\mathbf{Z} = \mathbf{W}\mathbf{E}$ and the columns of \mathbf{W}^{-1} are the common spatial patterns.

Although initially proposed in the field of BCI for Motor Imagery based BCIs, the same method has been successfully applied also for SSVEP based BCIs [55, 142]. In this case the two conditions A and B are given by the presence or the absence of a flickering stimulus and one filter is computed for the discrimination of each frequency.

In order to apply this method a “calibration session” is needed in order to compute the spatial filters and is organized as follow: various signal epochs are recorded

while the user is attending each stimulus; then the acquired signals are filtered for each of the stimulation frequencies by a narrow band pass filter centered on the stimulation frequency and a spatial filter is computed for each of the stimulation frequencies in order to increase the variance between the presence of the response to the particular frequency and its absence. To test for the presence of the response, the signal (previously band-pass filtered) is simply squared to obtain its power.

Once the spatial filters are computed for each frequency, the calibration phase is concluded and they can be used to train a classifier. For each of the stimulation frequencies the original signals are spatially filtered and band-pass filtered with the corresponding filters and their power computed. One classifier for each of the stimulation frequencies is then trained (the same data recorded to compute the spatial filters could be used) in order to distinguish the normal power, present in the EEG in that band, from the increased power arising in response to the flickering stimulation.

In the actual use of the BCI, the output of all the classifier is used to decide which of the stimulation frequencies is being attended by the user.

The main disadvantage of this method is given by the fact that it requires a calibration session where signals for the different conditions have to be acquired in order to compute the spatial filter. Moreover, across different session, the calibration has to be repeated since the position of the electrodes or their impedance could be slightly changed and thus also the optimal spatial filters to use may have to be changed accordingly.

MINIMUM ENERGY COMBINATION

The Minimum Energy Combination method [53] poses its foundation on a linear model of the EEG signal $y_i(t)$ read at the electrode i and measured as a voltage potential with respect to a reference electrode, while the subject is attending to a stimulus flickering at frequency f . In this model the signal is decomposed in three main components defined as follow:

$$y_i(t) = \sum_{k=1}^{N_h} a_{i,k} \sin(2\pi kft + \varphi_{i,k}) + \sum_j b_{i,j} z_j(t) + e_i(t) \quad (4.5)$$

The first component is the actual SSVEP response we want to detect, which is characterized by a set of sinusoids with frequency f and its k harmonics, each of which has an electrode specific amplitude $a_{i,k}$ and phase $\varphi_{i,k}$.

The second component of the model is a set of signals $z_j(t)$ that are unrelated to the SSVEP response and comprise concurrent brain activity and internal as external artifacts. These signals are present in all the electrodes i , scaled by the weighting factors $b_{i,j}$.

The last component $e_i(t)$ is a measurement noise component, specific to each electrode i .

In vector form, keeping the notion used in [53], for a time segment of N_t samples of the signal, sampled at a sampling frequency F_s , the model can be expressed as:

$$\mathbf{y}_i = \mathbf{X}\mathbf{a}_i + \mathbf{Z}\mathbf{b}_i + \mathbf{e}_i \quad (4.6)$$

where $\mathbf{y}_i = [y_i(1), \dots, y_i(N_t)]^\top$ is a $N_t \times 1$ vector and \mathbf{e}_i is a similar vector with noise. Meanwhile, the SSVEP model matrix \mathbf{X} is of size $N_t \times 2N_h$:

$$\mathbf{X} = [\mathbf{X}_1 \mathbf{X}_2 \dots \mathbf{X}_{N_h}] \quad (4.7)$$

where each sub-matrix \mathbf{X}_k contains a $\sin(2\pi kft)$ and a $\cos(2\pi kft)$ pair in its columns, while the $2N_h \times 1$ vector \mathbf{a}_i contains their respective amplitudes. The same holds for the \mathbf{Z} matrix, where the columns contains the noise signals and \mathbf{b}_i the respective weights. The model can be further generalized for multiple i electrodes as:

$$\mathbf{Y} = \mathbf{X}\mathbf{A} + \mathbf{Z}\mathbf{B} + \mathbf{E} \quad (4.8)$$

where $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_{N_y}]$ is a $N_t \times N_y$ matrix with the sampled signals from all the electrodes as columns and \mathbf{E} is a noise matrix constructed in the same way.

Regarding the noise, one of its components, that is for sure present as mentioned in Sec. 2.2, is an external artifact signal given by the power line frequency (e.g. 50 Hz in Europe and 60 Hz in USA), which for sure was not completely rejected by the bioamplifier common mode rejection system. This particular kind of noise can be modeled as \mathbf{Z}_p , a $N_t \times 2$ matrix containing a sine/cosine pair with the

power line frequency, thus a signal cleaned by this interference can be obtained as:

$$\mathbf{Y} \leftarrow \mathbf{Y} - \mathbf{Z}_p(\mathbf{Z}_p^\top \mathbf{Z}_p)^{-1} \mathbf{Z}_p^\top \mathbf{Y} \quad (4.9)$$

Given this model, the Minimum Energy Combination method has the goal to combine electrode signals into *channel signals* where the SSVEP response is magnified and the unrelated brain activity or noise is minimized. A *channel signal*, in this context, is a linear combination of electrode signals and thus a spatially filtered version of the original multi-electrodes recording. As for the CSP method, the goal is therefore to compute a spatial filter and consequently the multi-channel signal \mathbf{S} is obtained as a linear combination of the original recording \mathbf{Y} weighted by a spatial filter \mathbf{W} . Thus $\mathbf{S} = \mathbf{Y}\mathbf{W}$.

To compute \mathbf{W} using this method, the first step is to remove any potential SSVEP component from the recorded signals projecting them onto the orthogonal complement of the SSVEP model matrix \mathbf{X} , extracting $\tilde{\mathbf{Y}}$ which should contains all the unrelated brain activity and noise:

$$\tilde{\mathbf{Y}} = \mathbf{Y} - \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{Y} \approx \mathbf{Z}\mathbf{B} + \mathbf{E} \quad (4.10)$$

The next step is to compute a spatial filter able to minimize the resulting energy (or power) of the combination of electrode signals $\tilde{\mathbf{Y}}\hat{\mathbf{w}}$, resulting in the following optimization problem:

$$\min_{\hat{\mathbf{w}}} \|\tilde{\mathbf{Y}}\hat{\mathbf{w}}\|^2 = \min_{\hat{\mathbf{w}}} \hat{\mathbf{w}}^\top \tilde{\mathbf{Y}}^\top \tilde{\mathbf{Y}} \hat{\mathbf{w}} \quad (4.11)$$

The solution of the minimization is given by the the smallest eigenvector \mathbf{v}_1 and the energy of the resulting combination equals the smallest eigenvalue λ_1 . Choosing the eigenvectors as columns in the weight matrix \mathbf{W} , the obtained channel signals N_s in \mathbf{S} will be uncorrelated and will be ordered having an increasing uncorrelated activity and noise energy. Also the SSVEP response will be affected by the spatial filters, but it will be more easily detectable in the channels where the noise energy is lower. The weight matrix is then chosen as:

$$\mathbf{W} = \left(\frac{\mathbf{v}_1}{\sqrt{\lambda_1}} \dots \frac{\mathbf{v}_{N_s}}{\sqrt{\lambda_{N_s}}} \right) \quad (4.12)$$

To determine the number of output channels, N_s is determined as the smallest number for which:

$$\frac{\sum_{i=1}^{N_s}}{\sum_{j=1}^{N_y}} > 0.1 \quad (4.13)$$

thus N_s is chosen so as to discard as close to 90% of the uncorrelated brain activity and noise energy as possible.

Features describing the amplitude of the SSVEP response can then be computed from the spatially filtered signal (or signals) obtained from the original recordings of the electrodes.

4.3.2 SIGNAL TO NOISE RATIO

As mentioned at the beginning of this Section, one of the problems given by trivial methods as computing the FFT of the signals or computing the energy of narrow-band filtered version of the original signals, is that the SSVEP waveform power or energy is summed to the uncorrelated brain activity happening in the same frequency band.

To have a SSVEP response estimation less prone to influences given by other brain activities could significantly increase the obtainable ITR for SSVEP based BCIs applications. The ideal approach would be to succeed to divide the energy of the SSVEP waveform from the energy contributions of all the other brain activity in the same narrow band.

Several methods have been proposed to move towards this goal, for example using the FFT, a Signal-to-noise-ratio (SNR) could be computed dividing the power in the frequency of interest, with the power in the surrounding frequencies, where a similarly intense uncorrelated brain activity is supposed to be present, but where a SSVEP response should not [188].

Assuming f is the flickering frequency of a target (or one of its harmonics), the SSVEP response SNR in that frequency could be computed as:

$$SNR(f) = \frac{r |\mathcal{F}(f)|}{\sum_{k=1}^{r/2} |\mathcal{F}(f + k\Delta f)| + \sum_{k=1}^{r/2} |\mathcal{F}(f - k\Delta f)|} \quad (4.14)$$

where r is the even number of surrounding frequency to use, $\mathcal{F}(f)$ is the Fourier

coefficient of the signal at frequency f , while Δf is the Fourier transform frequency resolution (i.e. the distance in the frequency domain between adjacent bins).

Although this method proved to work, the assumption that the power of the underlying brain activity in the frequency of interest is roughly the same as in the adjacent frequency bands do not always holds. For example every subject have a peak in the natural brain activity in the alpha band named Intermediate Alpha Frequency (IAF) that could strongly interfere with the mentioned method.

A much better approach would be to remove the SSVEP waveform from the EEG signal in order to be able to measure the natural brain activity energy in the same narrow frequency band, to later use it to compute the SNR value for the original signal.

Indeed in [53], is proposed a statistic test to infer the SSVEP response intensity with respect to the noise, which have been adopted also in this work, that poses its bases exactly on this idea. The test, keeping the same notation as in Sec. 4.3.1, is defined as:

$$T = \frac{1}{N_s N_h} \sum_{l=1}^{N_s} \sum_{k=1}^{N_h} \frac{\hat{P}_{k,l}}{\hat{\sigma}_{k,l}^2} \quad (4.15)$$

where $\hat{P}_{k,l}$ is the estimated SSVEP power for the k -th harmonic frequency in channel signal \mathbf{s}_l and $\hat{\sigma}_{k,l}^2$ is an estimate of the noise and uncorrelated brain activity in the same frequency. In other words, the T statistic estimates how many time larger is the SSVEP response power compared to the case where no visual stimulus is present, averaging the SNRs ratios across N_h harmonics and N_s channel signals. The power $\hat{P}_{k,l}$ in the k -th harmonic frequency for the N_s channel signal is estimated as:

$$\hat{P}_{k,l} = \|\mathbf{X}_k^\top \mathbf{s}_l\|^2 \quad (4.16)$$

while, in order to avoid the need of calibration data acquired with no stimuli presentation and also to take into account the nonstationarity of the noise, the noise power $\hat{\sigma}_{k,l}^2$ is estimated on the same data segment used for the SSVEP detection, containing the SSVEP response. The SSVEP is therefore removed from the channel signals as shown in Eq. 4.10 to later fit an auto-regressive models $AR(p)$ of order

p to the channel signals and use the fitted models to interpolate the noise power in the SSVEP frequencies.

The $AR(p)$ models are fitted using the Wiener-Khinchin theorem for computing the autocovariance of each channel signal and then solving the Yule-Walker equations using a Levinson-Durbin recursion [53]. This yields the $AR(p)$ parameters $\alpha_1, \alpha_2, \dots, \alpha_p$ as well as an estimate of the variance $\hat{\sigma}^2$ of the white noise driving the $AR(p)$ process. Once fitted the model to the channel signal \mathbf{s}_i , the noise level estimated at the k -th harmonic is given by:

$$\hat{\sigma}_{k,l}^2 = \frac{\pi N_t}{4} \frac{\hat{\sigma}^2}{|1 + \sum_{j=1}^p \alpha_j \exp(-2\pi i j k f / F_s)|^2} \quad (4.17)$$

where N_t are the samples, k is the harmonic frequency number, f is the stimulation frequency in Hz, F_s is the sampling frequency in Hz and $i = \sqrt{-1}$.

The proposed test statistics have been widely adopted in various works and applied after the Minimum Energy Combination for spatial filtering, demonstrated to provide optimal results with respect to other methods [55].

Consequently in this work the SSVEP response intensity will be always evaluated using the Minimum Energy Combination followed by the T test statistic estimation, for multi-electrodes recordings, while using only T for single electrode recordings. The T test statistic in the following sections will be referred to as T index, in order to avoid a possible misinterpretation with the statistical *Student's t-test*, having a similar name.

In general, the use of SNR values as features, provides a further benefit regarding the different SSVEP response intensities over the three different frequency bands discussed in Sec. 4.2. In fact also the background uncorrelated brain activity has different intensities over these bands and its change in terms of power, with respect to the frequency band, is in some sense similar to the one of the SSVEP response. Consequently the use of SNR values leads to flatter curves [193] and also to a plateau where do not appear significant preferred frequencies over various subjects, as in the 11 Hz to 23 Hz range reported in [133]. The difference can indeed be appreciated comparing Fig. 4.2.2 with Fig. 4.3.2. This holds true in particular when the SNR is computed taking into account different harmonics frequencies

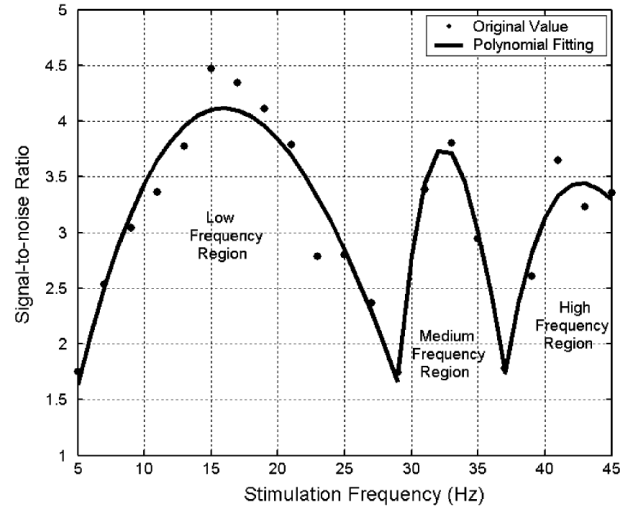


Figure 4.3.2: SNR of the SSVEP response with respect to the stimulation frequency, highlighting the three main components. Points are real data acquired from a real subject, while lines are polynomial interpolations. The same data depicted in Fig. 4.2.2 was used, but in this graph the SNR computed using a method similar to the one shown in Eq. 4.14 is plotted instead of the absolute power. Figure adapted from [193].

for the SSVEP response detection.

4.3.3 PHASE TAGGING

In the previous Chapters and Sections, it has always been implied that the different targets used for SSVEP based BCI applications need to flicker at different frequencies for the corresponding responses to be discriminated. Despite of this, it has been mentioned also the fact that the SSVEP response waveform is not only frequency coupled to the flickering stimuli frequency, but it is also phase coupled to the flickering phase of the stimuli.

Indeed it has been recently demonstrated that the phase lag of the SSVEP waveform, extracted from the EEG signals, is constant with respect to the phase of the flickering stimulus and that it can be used to detect which was the user's attended target [70, 209]. Therefore the *phase tagging* technique (in contrast to the *frequency tagging*), consists in the differentiation of the flickering targets not by their flicker-

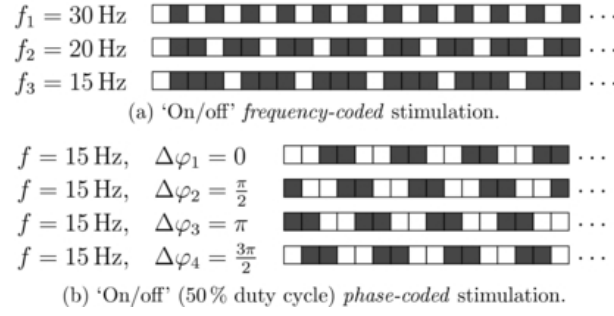


Figure 4.3.3: Schematic representation of the *frequency tagged* versus the *phase tagged* flickering sequence assuming as a stimulator device a regular 60 Hz computer monitor. Figure taken from [111].

ing frequency, but by their flickering phase.

To exploit the phase coupling to the stimulus could be used to increase the number of presentable targets given the same limited number of frequencies [78], or it could provide further information to filter the SSVEP response signal from the underlying EEG noise [90].

Unfortunately to extract the phase information alone is not enough to be able to detect which target was receiving the user's attention, since the lag between the first stimulus onset and the establishment of the steady-state (as was shown in Fig. 4.0.1) is unknown. Moreover, although the lag size is always in the same order of magnitude (80 ms to 160 ms), it is frequency and subject dependent. Consequently, what is needed to detect the attended target in a SSVEP based BCI implementing the *phase tagging* technique, is the phase difference between the SSVEP waveform and the flickering stimuli, plus the particular phase lag of the user for the used frequency.

The phase lag is known to be constant [78] for the same subject, given the same stimulus characteristics¹¹, consequently for targets flickering with different phases, the phase difference with respect to the SSVEP waveform, gives information about the attended target.

To obtain the phase difference between the flickering stimuli and the SSVEP response, the phase of the signal modulating the flickering has to be known and thus it has to be synchronously recorded with the EEG signals. In [208] a photo-

¹¹Anyhow, despite of the claimed phase constancy, it is still unclear how stable the difference could be over long periods of time [10].

diode has been used to record, along with the EEG, the modulating signal of the used flickering LED light, to later analyze it and compare its phase with the SSVEP waveform one, as shown schematically in Fig. 4.3.4.

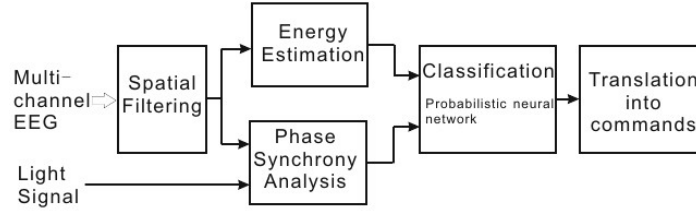


Figure 4.3.4: Scheme of an implementation of a *phase tagging* SSVEP based BCI, where the signal modulating the stimulus flickering is acquired by means of an external hardware device (a photodiode). Figure taken from [209]

To obtain a single signal from a multichannel recording, spatial filters, as the ones described in Sec. 4.3.1, could be applied before the phase analysis (as shown in Fig. 4.3.4), provided that are not introduced phase distortions¹². Otherwise, simpler approaches as a bipolar combination of two electrodes could be used [78], or more complex approaches as well, as the one proposed in [50], designed in particular for the phase analysis.

As proposed in [56], to determine the phase of a signal $x(t)$, the narrow-band frequency component corresponding to the frequency of interest f has to be isolated paying attention to use only linear-phase filters as FIR filters, otherwise the phase information would be corrupted. Only once a narrow-band signal $x(t)_f$ is obtained, it is possible to extract a meaningful phase information by means of the Hilbert transform as shown in Eq. 4.18.

Anyhow, since, as already mentioned, the interest is about the phase difference between the SSVEP waveform and the signal modulating the flickering of the target, the phase of both of them has to be computed and then compared. The Hilbert transform can be applied to both of the signals as shown in the continuous time formulation, respectively, in Eq. 4.18 and Eq. 4.19.

¹²Linear phase shifts are allowed since we are interested in the constancy of the difference with the flickering phase and not on the absolute phase.

$$A_{x_f}(t) = x_f(t) + j \mathcal{H}_{x_f}(t) = \rho_{x_f}(t) \cdot e^{j\theta_{x_f}t} \quad (4.18)$$

$$A_l(t) = l(t) + j \mathcal{H}_l(t) = \rho_l(t) \cdot e^{j\theta_l t} \quad (4.19)$$

where $\mathcal{H}_{x_f}(t)$ and $\mathcal{H}_l(t)$ are respectively the Hilbert transform of the narrow-band pass filtered signal $x_f(t)$, containing the recorded SSVEP waveform, and the Hilbert transform of the signal $l(t)$ modulating the flickering of the target. The instantaneous amplitude and phase computed for each signal are respectively ρ and θ , thus as proposed in [56], the phase difference for an epoch can be computed as the median value of the instantaneous phase difference $\delta_f(t)$ across the signal epoch as shown in Eq. 4.20.

$$\delta_f(t) = \theta_{x_f}(t) - \theta_l(t) \quad (4.20)$$

Others method of phase extraction could also be used, e.g. as proposed in [78], complex coefficients of the FFT could be computed and exploited to this aim.

The use of the *phase tagging* technique has proven to work for the discrimination of the gazed target. Various proof of concepts are available in the literature, but as stated also in [107], the use of this method for on-line classification has not been widely adopted yet, since it requires a very accurate real-time system able to keep synchronized the stimuli presentation, the EEG recording and the signal processing.

Despite of this, the phase information is statistically independent from the amplitude thus, to use both of them, demonstrated to lead to at least a factor of two improvement in the detection accuracy of a simple two-condition discrimination [90]. The use of mixed *phase* and *frequency tagging*, as recently shown in [78], seems indeed to allow for interesting ITRs (66.5 ± 18 bit/min estimated from an offline experiment) also in a more complex 15 targets discrimination task, using only 3 different frequencies.

The use of the phase information is therefore very interesting, in particular where few frequencies could be used and where accurate synchronization between the stimulus presentation and the EEG recording could be obtained. Most of the ex-

periments were accomplished using custom hardware as photodiodes, functions generators to drive LED lights, etc. but, as shown in [78] mixed *phase* and *frequency tagging* can be accomplished also on regular computer displays (although their experiment used an off-line classification).

4.4 PHOTSENSITIVE EPILEPSY

As already mentioned, electroencephalography is a non-invasive technique which do not entail medical risks for the subjects if performed with correctly working devices.

Despite of this, concerning SSVEP based BCIs, a possible hazard could be given by the presentation of SSVEP eliciting stimuli, since they may induce seizures in users predisposed to photosensitive epilepsy.

Photosensitive epilepsy (PSE) is a form of epilepsy in which seizures are triggered by visual stimuli that form patterns in time or space, such as flashing lights, bold, regular patterns, or regular moving patterns.

According to [47], between 4 and 9% of the population carries the risk of sensitivity to visually-induced seizures, which are induced by the physical characteristics of a visual stimulus and in particular photosensitivity seems to be greatest for flash frequencies between 9 Hz to 18 Hz, although nearly 50% of sensitive patients respond also to frequencies up to 50 Hz [4].

On the other side, according to [52], an abnormal EEG response to light or pattern stimulation, occurs in $\approx 0.3 - 3\%$ of the population, while the estimated prevalence of seizures from light stimuli is only ≈ 1 per 10,000 or 1 per 4,000 individuals aged 5 – 24 years. According to [52], the most provocative frequencies are in the range 15 Hz to 25 Hz and the red color also seems to be a factor. Red-cyan color combinations seems to be the most epileptogenic [4].

Unfortunately the most provocative frequency range for PSE is the same for the SSVEP eliciting stimuli, but at least, as a preventive measure the red color could be easily avoided.

Actually, the risk of photosensitive epilepsy is not related only to SSVEP eliciting stimuli, in fact seizures can be provoked by certain TV shows, movie screen images, video games, natural stimuli (e.g, sun on water), public displays, and many

other sources [52]. For example the “pocket monster” (Pokemon cartoon) incident on December 16, 1997 received a world-wide attention, as 685 children in Japan were treated for seizure symptoms after watching a television animated cartoon where large red frames alternated with blue frames at 12 Hz for several seconds [4, 52].

In fact, although guidelines for television broadcasting have been successfully implemented regarding allowed flicker frequencies, object sizes, alternating patterns and color compositions, most video games, but also pinball machines and other display devices, are supplied with warning labels informing about the risk of seizures for photosensitive users.

To reduce the risk of PSE, for all the performed experiment the red color was not used for flickering targets and moreover all the subjects were informed of the risk and asked if they ever had seizures or if they were aware of being predisposed to PSE.

5

Hardware and Software tools

In this chapter will be described the hardware available to perform the experiments and will be illustrated the utilized software tools.

In order to use existing state-of-the-art techniques in conjunction with the available hardware, custom software development has been necessary and in this chapter will be discussed as well.

Thanks to the available information about the SSVEP response illustrated in Chap. 4, a state-of-the-art software able to provide reliable stimulations for SSVEP elicitation has been developed from scratch.

The developed software, apart from the stimuli presentation reliability, provides also the needed flexibility to perform experiments in particular contexts, such as stereoscopic displays, and to change various stimuli parameters, as will be detailed hereby as well. Moreover it implements a state-of-the-art software synchronization mechanism with the EEG acquisition software.

Eventually, using a pre-existing state-of-the-art software framework for the implementation of custom BCI systems and a custom implemented signal processing

pipeline, an actual SSVEP based BCI system has been implemented.

The obtained system is meant to be a step beyond the state-of-the-art, combining together the best performing signal processing methods, the latest pre-existing data acquisition and managing software for BCIs and the most precise and flexible stimuli presentation software obtainable.

5.1 ACQUISITION DEVICES

At the Eidomatic Laboratory of the Department of Computer Science of the University of Milan are available two different EEG acquisition devices, a commodity device with a single dry electrode: the Mindset by NeuroSky Inc. detailed in Sec. 5.1.1 and a professional low-end gel-based passive electrode device with 4 EEG channels named multipurpose g.MOBilab+ by g.tec medical engineering GmbH detailed in Sec. 5.1.2.

5.1.1 NEUROSKY MINDSET

The Mindset, often referred as a “toy EEG” is a very simple and commodity EEG acquisition device, looking like regular headphones, equipped with a single dry electrode as depicted in Fig. 5.1.1.



Figure 5.1.1: The Mindset acquisition device produced by NeuroSky Inc. Figure taken from <http://www.designboom.com/>.

The single electrode is designed to be positioned on the forehead, roughly at the Fp1 position with respect to the 10-20 system described in Sec. 2.2.1. It acquires the EEG signal band-pass filtered between 3 Hz and 100 Hz at a sampling rate of 512 Hz, digitizing it at 12 bit. It can be connected to a computer for data acquisition using a Bluetooth connection. It incorporates a notch filter to remove power-line artifacts and implements proprietary algorithms for further signal cleaning and feature extraction.

In addition to the single acquisition electrode, the MindSat has also three other contacts to be positioned over the left ear of the subject, which are used as ground and reference electrodes.

In particular, apart from the raw filtered EEG signal, it provides also proprietary dimensionless features representing the power strength in the clinical frequency bands described in Sec. 2.1.4 and also two 1 Hz sampled signals called *e-Sense Attention* and *e-Sense Meditation* values.

The *Attention* and *Meditation* values are computed thanks to a proprietary algorithm and very few information are available about their actual meaning. In the manufacturer intentions, the subject wearing this device should be able to learn to control these two values in order to be able to use *Active* BCI applications after some training.

The device is shipped with a software bundle including a self-person maze computer game where the Mindset input is used as a secondary input (to lift, pull, burn, etc. game objects) in addition to the ordinary keyboard arrow keys ¹.

A device driver for the MindSet is already available in the OpenVibe software framework which will be introduced in Sec. 5.3 and it permits to extract both the raw signal and the proprietary computed values. Interestingly the MindSet has an automatic on-line check to detect the contact quality between the electrodes (single electrode plus ground and reference). The contact check reading can be acquired by software too, but instead of an impedance value, it returns a SNR (signal to noise ratio) between what a proprietary algorithm considers as the EEG signal and what it identifies as noise.

¹<http://store.neurosky.com/products/the-adventures-of-neuroboy-bci-technology>



Figure 5.1.2: g.MOBilab+ multipurpose biosignal acquisition system manufactured by g.tec medical engineering GmbH. Figure adapted from [64].

5.1.2 G.TEC G.MOBILAB+ MULTIPURPOSE

The g.MOBilab+ multipurpose version is a portable biosignal acquisition device utilizable to acquire EEG, electrocardiogram (ECG), electrooculogram (EOG) and electromyogram (EMG) bio-signals. It is equipped with low-noise biosignal amplifiers and a 16 bit analog to digital converter sampling at 256 Hz. It can be connected to a regular computer through a Bluetooth connection for data acquisition.

It has 8 channels as the regular version, but only four of them can be used to acquire EEG signals, since to be multipurpose, different signal amplitudes have to be taken into consideration and different amplifiers with different gains have to be used (e.g. EEG signals commonly have an amplitude of several microvolts while ECG signals have an amplitude of few millivolts). In particular the first two channels can be used only for EEG signals acquisition, while channel 3 and 4 can be used both for EEG and EOG. Channels 5 and 6 can be used for ECG/EMG signal acquisition, while channel 7 and 8 can be used as generic analog inputs.

The multipurpose nature of this device, in particular regarding the first 4 chan-

Channel	Sensitivity	High pass	Low pass
1	$\pm 500 \mu V$	0.5 Hz	100 Hz
2	$\pm 500 \mu V$	0.5 Hz	100 Hz
3	$\pm 2000 \mu V$	0.01 Hz	100 Hz
4	$\pm 2000 \mu V$	0.01 Hz	100 Hz

Table 5.1.1: g.Tec g.MOBilab+ multipurpose version channel sensitivities and hardware filters specifications.

nels, which will be used in this work for EEG acquisition, is worth to be analyzed in more details since differences between channels 1 and 2 exist with respect to channels 3 and 4. In particular different hardware channel sensitivity and filters are used as detailed in Tab. 5.1.1.

The different sensitivities, in order to acquire EEG signals using all of the first 4 channels, are compensated through the software at the driver level, multiplying the digits acquired from the device by a conversion factor, to obtain a value expressed in microvolts. Due to this operation, since the digital to analog converter is using its 16 bit to span the whole sensitivity range for every channel, will lead to an equivalent digitization of 14 bit for the channel 3 and 4, assuming that a linear quantization is used, since not explicitly stated otherwise.

This is worth to notice since in the acquired signals from channels 3 and 4 are pretty evident higher low amplitude fluctuations (due to the different high-pass filter), but also, looking at the signal PSD, a stronger power for higher frequencies, probably due to an higher quantization noise in these channels, caused by the different sensitivity.

5.2 STIMULI PRESENTATION DEVICES

To present the flickering stimuli, for SSVEP based BCIs applications, different devices could be used as discussed in Sec. 4.1.3. In this Section the devices available in the context of this research work will be presented and their main specifications will be described.

In compliance with the observations given in Sec. 4.2.3 and Sec. 4.1.3 a descrip-

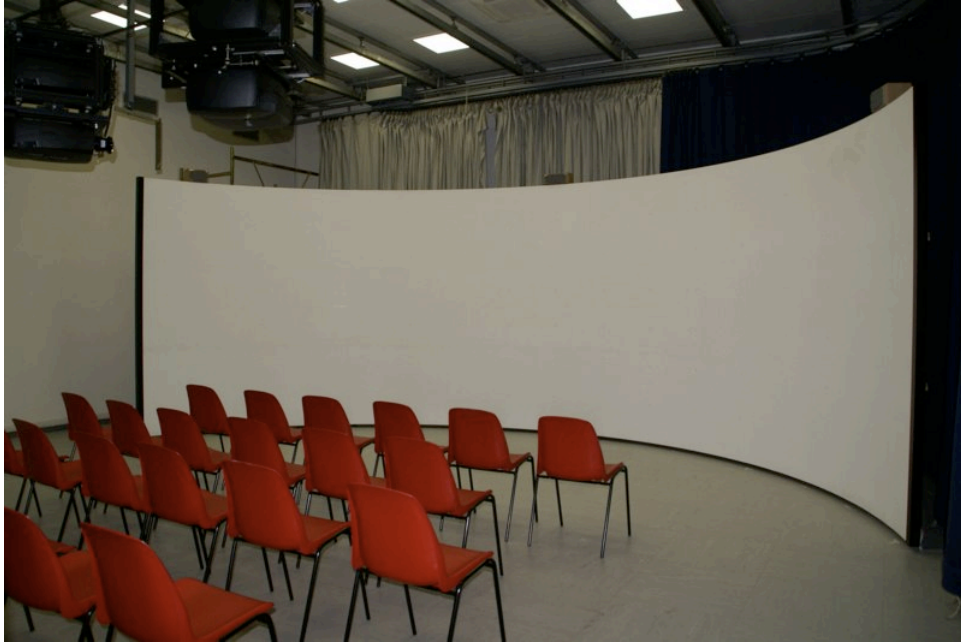


Figure 5.2.1: Virtual Theater installation at Università degli Studi di Milano.

tion of the devices spectral characterization will be given as well.

5.2.1 VIRTUAL THEATER

The Virtual Theater of the University of Milan, shown in Fig. 5.2.1, is driven by four BARCO Sim5 Plus projectors mounted on two metallic chassis, organized in two horizontal couples, each couple projecting a 2416×1050 image, covering a field of view of 120° horizontally and of 90° vertically, from an observation distance of 3 m.

The projection screen is an highly reflective wide curved semi-cylindrical screen, having an height of 2.70 m, with a radius of 3 m and an arc length of 8 m.

The four projectors can be used at a maximal refresh frequency of 60 Hz and they are able to provide a stereoscopic visualization mode based on the INFITEC color filters [79], requiring the use of specific passive glasses for the users.

The light spectrum of the three RGB components separated and summed (the white light), measured after the screen reflection, without the INFITEC filters are reported in Fig. 5.2.2.

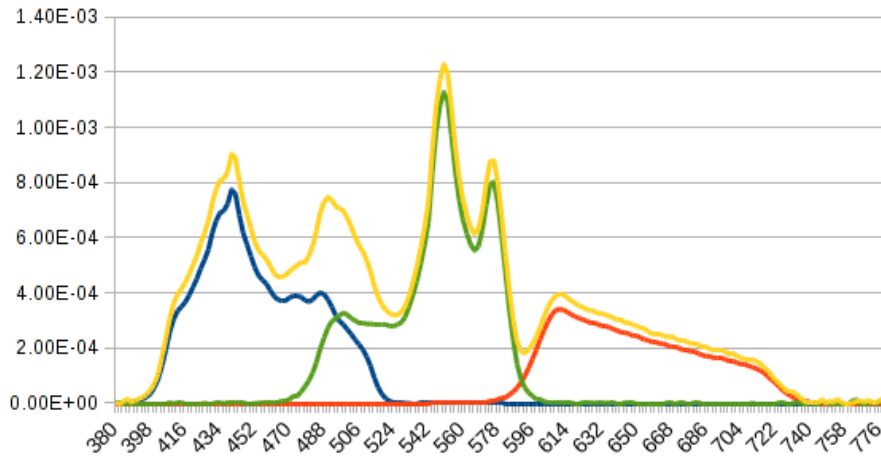


Figure 5.2.2: Light spectrum of the three single RGB components plus the white of the Virtual Theater after the screen reflection. Dataset acquired during the experiments described in [54]. On the x-axis the wavelength in nm while on the y-axis the spectral radiance in $\text{W sr}^{-1} \text{m}^{-3}$.

In the context of this research work, the aforementioned Virtual Theater was not used as a stereoscopic device, since the INFITEC filters modifying the light spectrum content differently for the two eye views [79] would induce different SSVEP responses for the two eye visual channels in the brain. This phenomena would probably introduce yet unknown responses combinations, which although worth of further investigations, were not considered in the scope of this work.

5.2.2 ASUS VG278H

The ASUS VG278 is a modern commodity computer monitor integrating a stereoscopic modality based on the NVIDIA 3D Vision technology ². It features a 27" screen, with a 1920×1080 pixel resolution resulting in a 16 : 9 aspect ratio. The panel is based on a twisted nematic (TN) display back lit by LED lights, with a maximal refresh rate of 120 Hz, when used as a regular monoscopic monitor. In Fig. 5.2.3 is reported its light spectrum (for distinct RGB components and their sum) measured using a GretagMacbeth Eye-One spectrophotometer.

²<http://www.nvidia.com/object/3d-vision-main.html>

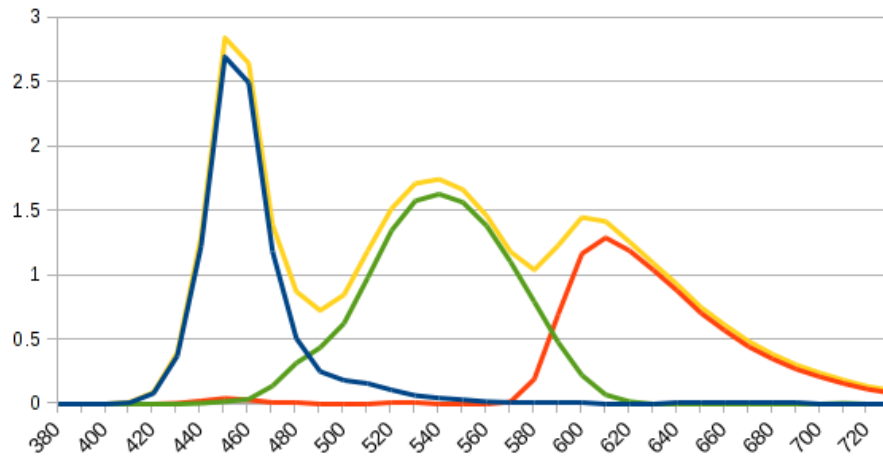


Figure 5.2.3: Light spectrum of the three single RGB components plus the white of the ASUS VG278 monitor. Dataset acquired with a GretagMacbeth Eye-One spectrophotometer. On the x-axis the wavelength in nm while on the y-axis the luminance normalized spectral radiance.

When used as a stereoscopic monitor, thanks to an embedded IR emitter, it is able to drive liquid crystal (LC) shutter glasses in order to alternatively suppress the light entering the user's left and right eye. In this operation mode, the monitor display alternatively left and right eye views which are respectively let enter the user's left and right eye by the glasses. The monitor can therefore reach a maximum of 60 Hz refresh rate for each of the eye views.

The stereo driver proprietary software can perform automatic stereoscopic conversion by using the 3D models submitted by the application and rendering two stereoscopic views instead of the standard mono view.

Fortunately, with respect to the experiments described in Sec. 6.2, the standard *quad-buffering* mode (which will be detailed in Sec. 5.4.2) can be used as well, allowing developers to control the rendering, avoiding the automatic mode of the driver, in order to just render independently the scenes to the left and right frame buffers.

5.2.3 DELL P2210F

The DELL P2210f is a modern commodity regular computer monitor. It features a 22" screen, with a 1680×1050 pixel resolution resulting in a 16 : 10 aspect ratio. The panel is based on a twisted nematic (TN) display back lit by cold-cathode fluorescent lamps (CCFL) lights, with a nominal refresh rate of 60 Hz. In Fig. 5.2.4 is reported its light spectrum (for distinct RGB components and their sum) measured using a GretagMacbeth Eye-One spectrophotometer.

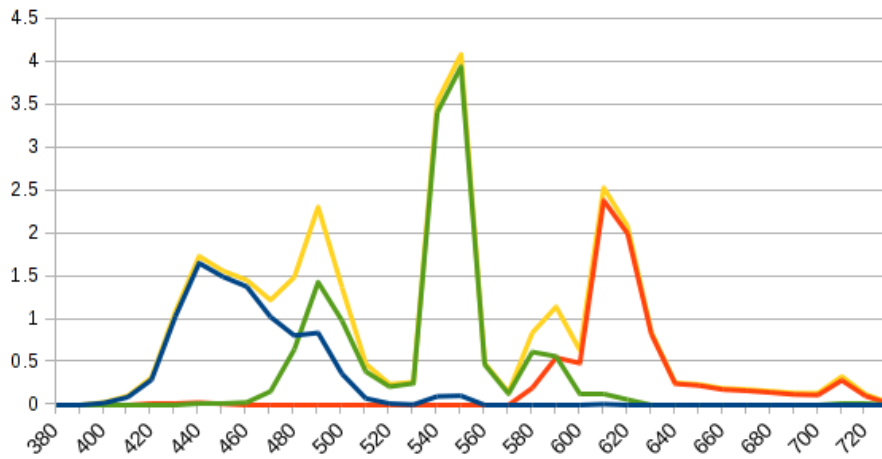


Figure 5.2.4: Light spectrum of the three single RGB components plus the white of the DELL P2210f monitor. Dataset acquired with a GretagMacbeth Eye-One spectrophotometer. On the x-axis the wavelength in nm while on the y-axis the luminance normalized spectral radiance.

5.3 THE OPENVIBE SOFTWARE

Although various BCI implementations exist, most of them were programmed within single research groups to fulfill specific requirements and were not meant for a wider adoption. General tools for off-line and on-line EEG analysis are also available, but comprehensive frameworks offering enough flexibility for BCI implementations are limited [17]. The most important frameworks which could be

considered as comprehensive set of tools for generic BCI implementations are BioSig, BCI2000, BCI++ and OpenVibe [17].

Interestingly OpenVibe [163], developed at INRIA³, is distributed under a Free Software license, it is multi-platform and has been developed with the aim of producing a general framework for BCI implementations for a wide diffusion. Moreover it has been designed having in mind in particular the context of VR environments, where it has been used in different research works [109].

For these reasons also in this work, for all the performed experiments, the EEG signals have been acquired using the OpenVibe software [163] and also the events/stimuli triggering have been handled within this environment.

OpenVibe major strengths, apart from the openness of its source code, which created a growing community around it, are: its modularity, since it is made of different applications and blocks; its flexibility in terms of implementable configurations and its capability of communicating with other software tools thanks to various network protocols and file formats.

OpenVibe is divided in two main applications, the *OpenVibe Acquisition Server* and the *OpenVibe Designer*. The *Acquisition Server* is meant to be connected to the EEG device in order to manage the data stream acquisition and it is provided with a set of different drivers able to manage different signal acquisition hardware devices. A more detailed description can be found in Sec. 5.3.1.

On the other side, the *Designer* is a graphical tool utilizable to catch a stream of EEG data and triggering signals, from one or more *Acquisition Servers*, in order to implement custom data processing, interconnecting in a graphical data-flow different processing boxes. Every graphical box actually contain a library for data processing or feature extraction, triggering signals handling, etc. A more detailed description of the *Designer* and the available boxes can be found in Sec. 5.3.2.

A scheme for a generic BCI implementation using the OpenVibe framework, highlighting to closed-loop control for a generic VR application is shown in Fig. 5.3.1

³<http://www.inria.fr/en/>

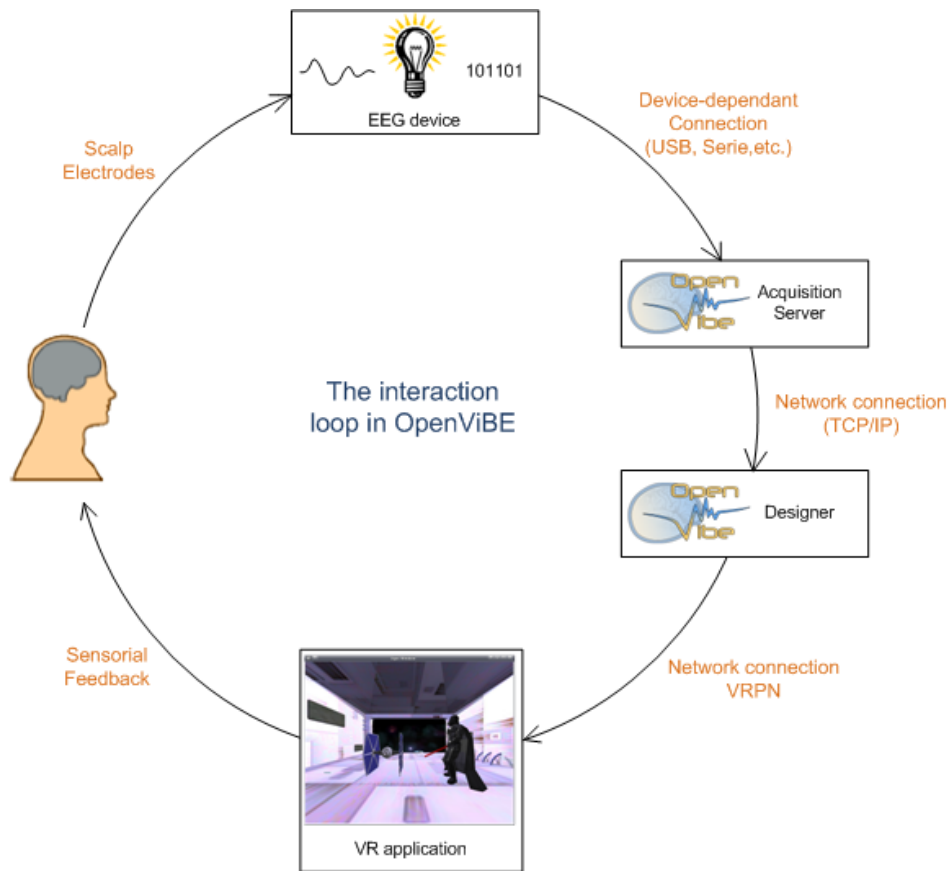


Figure 5.3.1: Scheme of a generic BCI using the OpenVibe framework to control a generic VR application. The closed-loop interaction is highlighted in particular, as the kind of data transmitted in each part of the control loop. Figure courtesy of <http://openvibe.inria.fr/> .

5.3.1 THE ACQUISITION SERVER

The *Acquisition Server* provides a generic interface to various kinds of acquisition devices, (e.g. EEG or MEG devices). Such an abstraction allows the user to create hardware independent applications, thanks to the use of a generic acquisition box in the implemented *Designer* scenario, as the one shown in Fig. 5.3.2.

The generic acquisition box receives data via the network from the *Acquisition Server*, which is actually connected to the hardware and could provide generic data streams of signals and triggering events.

The way the acquisition server gets connected to the device mostly depends on the hardware manufacturer's policy. Some devices use standard communica-

tion protocols over the network, over serial/USB ports or over bluetooth with well documented data formats, while some others may implement proprietary protocols or undocumented data formats requiring the use of specific SDK (Software Development Kits) or shared libraries.

The role of the *Acquisition Server* is especially to keep all of these hardware peculiarities out from the actual data handling and processing part which in turn is implemented in the *Designer*. Thanks to the *Acquisition Server*, the *Designer* is fed with a standard data stream always with the same device-independent data format.

A recently added feature in the *Acquisition Server* is the capability of merging with the acquired signal also software triggering signals coming from external applications. Triggering signals are often used in the context of BCI and neuroscience experiments, since it is often needed to store within the EEG signals some timestamps denoting a particular event as for example the initiation of a visual stimulus. In the past, hardware triggering was commonly used and professional devices still often incorporate generic digital inputs for triggering purposes, able to receive information from custom hardware. Nevertheless, in the case of VR applications it is particularly inconvenient to instruct serial/parallel ports to generate hardware triggers via software to later digitize them again. Moreover commodity devices do not implement any port for external signaling.

Before this feature was introduced I faced this problem implementing a box for the *Designer* able to acquire software triggers through an UDP socket, but this approach demonstrated to work with time alignment errors in the order of several tens of milliseconds, which for some applications may be acceptable, but for others it is not (e.g. for SSVEP phase analysis).

Software triggering is now implemented in the *Acquisition Server* using the Boost Inter-Process Communication (IPC) library, providing the fastest communication possible between two applications. This approach has the only limitation of requiring both the applications (VR application and the *Acquisition Server*) to run on the same machine; despite of this, in my tests it proved to reliably provide software triggering with time alignment errors lower than the EEG device sampling period.

5.3.2 THE DESIGNER

The *Designer* is mainly dedicated to the BCI application author and enables him/her to build complete scenarios based on existing software modules using a dedicated simple Graphical User Interface (GUI) as shown in Fig. 5.3.2.

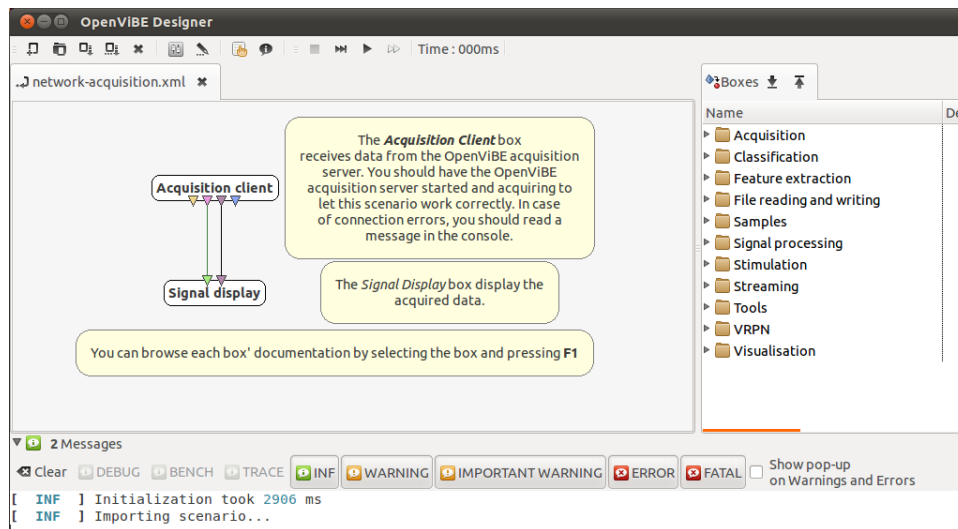


Figure 5.3.2: Screenshot of the OpenVibe *Designer* with a simple scenario to just acquire a data stream from the *Acquisition Server* and display it. Figure adapted from <http://openvibe.inria.fr/> .

The author has access to a list of existing modules named boxes which could be drag and drop in the scenario window. Each module appears as a rectangular box with inputs on top and outputs at the bottom. Boxes are manually connectable through their inputs/outputs and different kind of streams exist. The two most important streams are *signal* and *stimulation*, where with the word *stimulation* are meant triggering signals, probably since they are commonly associated to stimuli events. *Signal* streams on their turn could carry different kind of data, from matrix containing signals samples per channels to feature vectors containing features extracted from the original signals.

An embedded player engine allows the author to test and debug a scenario in real time. In doing so, the author can receive continuous feedback on the boxes status and processing times. Such feedback could be useful also to balance the computational load.

The *box* is the key component of the platform; it consists of an elementary component in charge of a fraction of the whole processing pipeline. Each box can be notified on clock ticks and upon input data arrival in order to activate and execute a generic code segment. The characteristics and constraints that are common to all boxes include reasonable granularity to allow quick software components rearrangement.

Different boxes are available implementing commonly used algorithms in the field of BCIs, but also generic ones are available, able to send data to custom scripts and to receive the results back. Scripts could be written in the Lua programming language⁴, but a recently introduced experimental box allow also the use of scripts written in the MATLAB language.

COMMUNICATION WITH THE EXTERNAL APPLICATIONS

The OpenVibe framework is meant to be used to implement a generic BCI controlling an external application which is commonly not directly implemented in the *Designer*. A typical example would be to control a pre-existing VR environment or Computer Game, as shown in Fig. 5.3.1.

One of the most common used protocol to let VR environments to receive inputs from generic devices is, as already mentioned, the VRPN protocol. OpenVibe implements indeed various VRPN boxes in the *Designer*, providing both VRPN client and server services.

Taking into account the *software tagging* feature recently introduced, as well as all the mentioned components of the OpenVibe framework, a generic connection with an external application (as could be a VR environment or Computer Game) is sketched in Fig. 5.3.3.

Using custom boxes, further communication protocols could be implemented as UDP/TCP sockets, moreover, external applications could be launched thanks to a box able to run a generic executable file.

⁴<http://www.lua.org/>

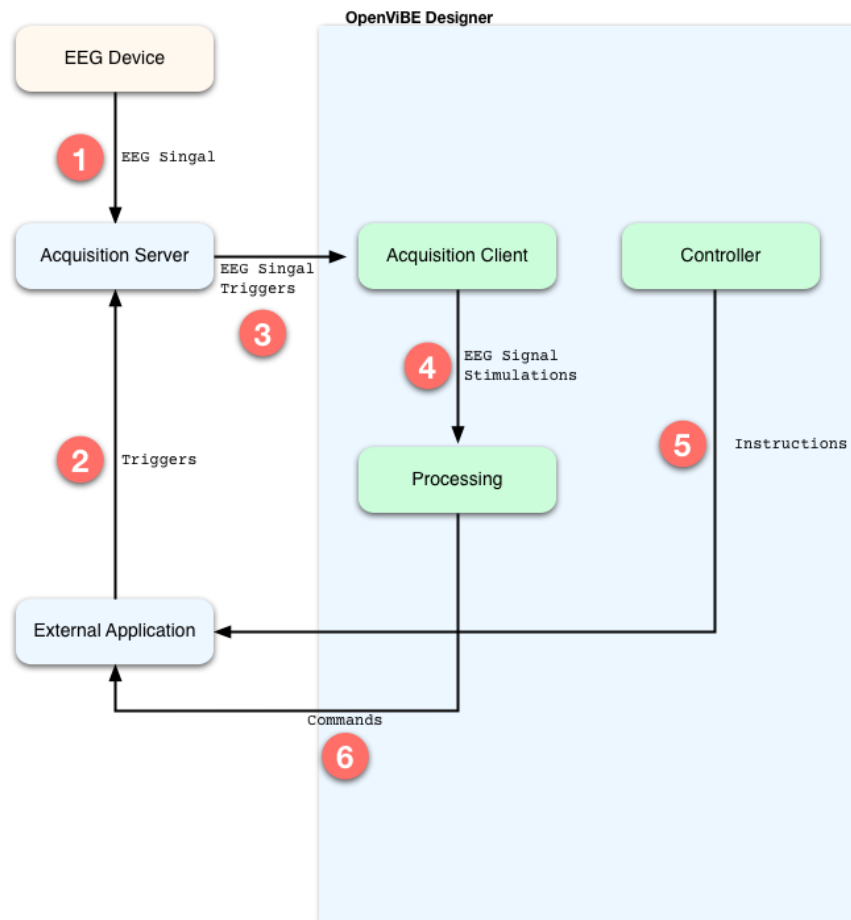


Figure 5.3.3: OpenVibe software tagging schema. (1) The OpenViBE Acquisition Server acquires signals from the EEG device. (2) At the same time the *External Application* sends triggers to the *Acquisition Server*. (3) The *Acquisition Server* combines signal from the EEG and triggers from the *External Application* into one stream, triggers are named *Stimulations*. (4) The *Acquisition Client* box will pass the signals to the signal processing chain. (5) An optional controller box can give commands to the *External Application* via VRPN. (6) The processing chain will give commands to the *External Application*. Figure courtesy of <http://openvibe.inria.fr/> .

5.3.3 G.MOBI LAB+ DRIVER DEVELOPMENT

In order to use the g.Tec g.MOBIlab+ multipurpose acquisition device within the OpenVibe acquisition server a device driver had to be developed since it was missing from the available OpenVibe's devices.

A GNU/Linux only driver for the EEG-only version of the g.MOBIlab+ device developed by Lucie Daubigney from Supelec Metz was available within the OpenVibe community and was used as the starting source code.

Initial development of the driver was started by Andrea Villa, an undergraduate student at the University of Milan who succeeded to develop a Windows only version capable of acquiring a single electrode signal from the g.MOBIlab+ device.

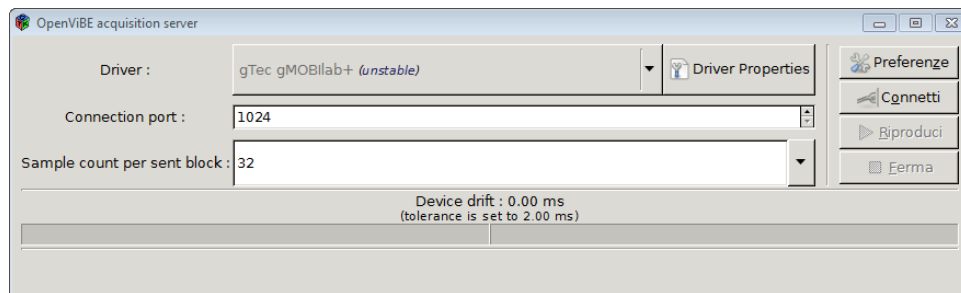


Figure 5.3.4: The graphical user interface of the OpenVibe *Acquisition Server* integrating the developed g.MOBIlab+ driver.

The final development was accomplished by me, enabling the driver to use all of the electrodes and providing a multi-platform version runnable both on Windows and GNU/Linux operating systems, able to work with both the multipurpose and EEG-only versions of the g.MOBIlab+ device. In Fig. 5.3.4 is shown the OpenVibe *Acquisition Server* ready to acquire data from the g.MOBIlab+.

Moreover, as shown in Fig. 5.3.5 the *Acquisition Server* graphical user interface has been modified to integrate particular options provided by this device as the “test mode”, to check the acquisition without actually amplifying the electrode signals, which is a very handy tool to avoid possible damages to the electronics while doing test acquisitions with electrodes not connected to the subject’s scalp.

The driver has been later contributed to the OpenVibe community under the

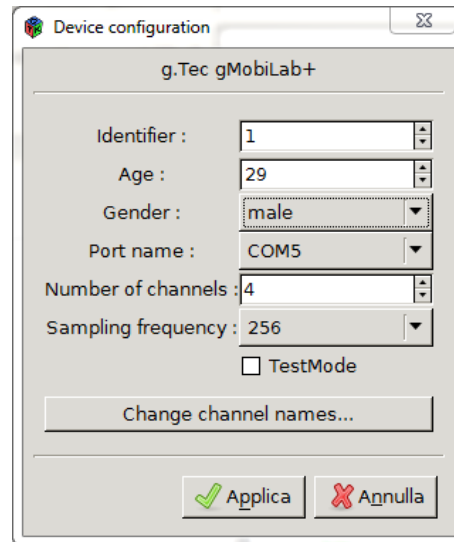


Figure 5.3.5: The device configuration window of the OpenVibe *Acquisition Server* integrating the developed g.MOBilab+ driver.

GNU GPL v3.0 license ⁵.

5.3.4 MINIMUM ENERGY COMBINATION BOX DEVELOPMENT

Within the OpenVibe framework are available several scenarios, to be loaded in the *Designer*, implementing the most common BCI paradigms as Motor Imagery, P300 and SSVEP based BCIs.

The already available scenario implementing the SSVEP based BCI paradigm is actually composed of five different scenarios to be loaded in the *Designer* and to be run in the right sequence ⁶.

The first scenario is used just to set some configuration files with various parameters as the flickering frequencies, target sizes, colors, epochs length, etc.

A second scenario is used to acquire training data; it executes the OpenVibe stimuli presentation software, showing to the subject all the targets, plus a marker on the target to be attended, in order to record labeled epochs to be later used for classifier training.

A third scenario is then used to replay the training data previously acquired in

⁵<http://openvibe.inria.fr/forum/viewtopic.php?f=14&t=623&start=15>

⁶<http://openvibe.inria.fr/steady-state-visual-evoked-potentials/>

order to compute spatial filters using the CSP method described in Sec. 4.3.1.

A fourth scenario, applying the previously computed spatial filters to the incoming signals, trains one classifier for each of the stimulation frequencies in order to detect a corresponding SSVEP response.

Eventually, the last scenario, using the previously computed spatial filters and classifiers, implements the actual on-line BCI.

The already available SSVEP based BCI implementation is therefore based on the CSP method and to detect the presence of a SSVEP response is exploiting as features the narrow-band power estimations of the spatially filtered EEG signals.

As detailed in Sec. 4.3.1 and Sec. 4.3.2, more sophisticated method exists, which proved to provide better performances and moreover to be able to shorten the training time requested.

Consequently I implemented, with the help of Mariangela Littini, an undergraduate student, the SSVEP detection procedure presented in Sec. 4.3.2 in a custom OpenVibe box able to compute the T index for EEG data acquired on-line within the OpenVibe framework, using the Minimum Energy Combination method detailed in Sec. 4.3.1.

The feature extraction algorithm was implemented as a MATLAB script and I was able to include it in a custom OpenVibe box using a recently introduced experimental box⁷ able to call a MATLAB/Octave function sending data chunks for generic signal processing and to read results back.

The implemented box has been designed in order to be fully customizable from the OpenVibe *Designer* window, where can be configured the frequencies to be used as well as the number of harmonics to be evaluated.

This box accepts as an input an epoched data stream with custom epoch length and returns as an output a feature vector of T indexes, having a dimension corresponding to the number of configured frequencies, e.g. for a BCI with three targets, the feature vector would be: $\langle T_{f_1}, T_{f_2}, T_{f_3} \rangle$.

⁷<http://openvibe.inria.fr/tutorial-using-matlab-with-openvibe/>

5.4 STIMULI PRESENTATION SOFTWARE DEVELOPMENT

To present reliable flickering stimuli to elicit SSVEP responses using ordinary displays is known to be challenging because of the strict timing constraints to meet, in order to obtain a strong response. Moreover, apart from the timing issue, the stimulus has to be carefully controlled in order to fine tune its properties, as its duty cycle, color, contrast, etc. as detailed in Chap. 4.

The same constraints hold also when using other kind of stimulation devices, as LED lights, but in that case dedicated hardware is commonly used, as micro-controller chips or waveforms generators, providing an highly reliable frequency and modulation control.

In fact, also during my research work, for preliminary testing, I built a LED stimulator controlled by an Arduino Due board, shown in Fig. 5.4.1, similar to the one proposed in [180], able to reliably flicker high power LED lights at any frequency in the range useful for eliciting a SSVEP response in a subject gazing at it.

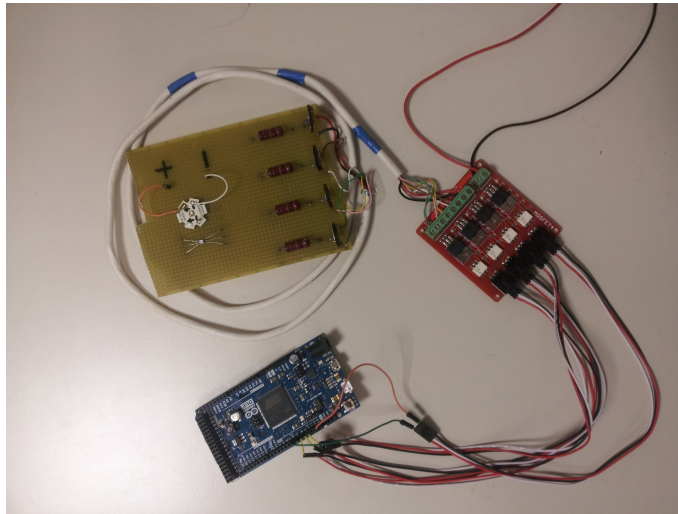


Figure 5.4.1: My custom built LED stimulator prototype. The blue board is an Arduino Due board able to control four digital pins used to open and close four power FET gates hosted on the red board which drive current to the 3 channels of an RGB LED and to an interchangeable monochromatic power LED, hosted on the yellow board.

Anyhow for the use of SSVEP based BCIs in the context of VE, the use of or-

dinary displaying devices commonly used for VE visualization, has to be foster, to present integrated SSVEP eliciting stimuli, in order to maintain immersivity and presence [109].

One of the most used and popular software tool used to provide reliable and highly customizable visual stimuli using ordinary computer screens is the Psychophysics Toolbox (Psychtoolbox) [16]. It is a free ⁸ set of Matlab and GNU/Octave functions for vision research, able to easily synthesize and show accurately controlled visual and auditory stimuli.

The Psychtoolbox has been already used to provide SSVEP eliciting stimuli for BCI applications, but although being an handy tool to provide various kind of stimuli for different experimental scenarios, it is not easily integrable within VE graphical engines. The Matlab and GNU/Octave language is indeed an interpreted scripting language which was not intended to be used for computer graphics.

From the Psychtoolbox Version 3 (PTB-3), its Matlab extensions (written in C) were rewritten [89] in order to be more modular and to use OpenGL (Open Graphics Library) [172]. The Psychtoolbox is consequently a valuable tool able to grant to Matlab and GNU/Octave users the possibility to display highly controllable visual stimuli, but would not be the best choice for the integration of flickering stimuli in VEs, which are commonly already programmed using compiled languages as C/C++ and graphics libraries such as OpenGL.

Therefore, in this Section the OpenVibe solution to provide controllable flickering stimuli integrated in VEs will be presented.

Moreover, will be addressed in details a custom developed solution developed in order to provide a generic highly precise and highly customizable tool able to be easily integrated in existing VEs.

5.4.1 OPENVIBE SOLUTION

In the OpenVibe software, described in Sec. 5.3, is available a scenario implementing a SSVEP based BCI, using the CSP algorithm described in Sec 4.3.1, which exploits for the flickering stimuli presentation an ad-hoc application written using the OGRE Environment.

⁸Mostly covered by the MIT license or a MIT compatible license.

OGRE (Object-Oriented Graphics Rendering Engine) is a scene-oriented, real-time, flexible 3D rendering engine written in C++, designed to make it easier and intuitive for developers to produce applications utilizing hardware-accelerated 3D graphics.

This application purpose is to present flickering targets on an ordinary computer monitors, as discussed in Sec. 4.1.3, displaying patterns as the ones sketched in Fig. 5.4.2, where every box represent a screen frame and white/black colors relate to the on/off state of each target at the wanted frequency.

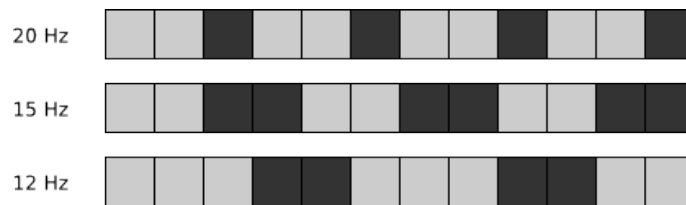


Figure 5.4.2: Stimulation pattern for SSVEP stimuli presentation using a regular 60Hz computer displays. Figure courtesy of <http://openvibe.inria.fr/>.

Using this application provided with OpenVibe, evident and clearly perceptible flickering frequency instabilities have been noticed. Different hardware (computers and displays) and operating systems have been tried in order to diagnose the cause of this malfunctioning. Despite of this, although with different intensities, a stable flickering at least at a naked eye inspection could not be obtained in any case.

The same problem has been noticed by other users in the OpenVibe community and has been reported in the project forum ⁹. Also using the stimulator in full-screen mode as suggested did not solved the problem, thus in order to be able to control with finer details how the synchronization is managed, I developed a custom solution without using the higher abstraction provided by the OGRE environment over the lower level libraries.

⁹<http://openvibe.inria.fr/forum/viewtopic.php?f=17&t=591&p=8466>

5.4.2 CUSTOM SOLUTION

It is of great interest for this application, not only to have access to precise timers, but in particular to have access to the synchronization mechanism with the screen refresh, as demonstrated in [30].

The stimuli presentation software I developed is written in C++ language and based on OpenGL [172] which is widely used in the context of computer graphics.

Most of the VR environment implementations are based on the OpenGL library, although higher level programming environments are often used to provide to developers an easier abstraction layer, as OGRE for example. Unfortunately higher abstraction layers use to hide lower level details that are commonly not needed, but whose in the particular case of the presentation of stimuli to elicit SSVEP responses are useful to have an higher control over the synchronization mechanism.

The OpenGL library is a cross-language, multi-platform API for rendering 2D and 3D computer graphics and is typically used to interact with a GPU, to achieve hardware-accelerated rendering. In practice the OpenGL library permits to easily program Computer Graphics (CG) and thus in simple words to draw 2D or 3D shapes on a screen. Changing the positions of the scene objects or the position of the virtual camera for every frame, CG animations could be generated.

The OpenGL library is quite complex and permits to run different kind of contexts over different kinds of hardware, anyhow the three main features of the OpenGL library which have been exploited for this implementation are the double-buffering capability, the texture caching and in particular the automatic synchronization of the buffers swap with the screen refresh. In the following description will be mentioned only the used OpenGL contexts and will be omitted all the different options which are not strictly of interest for the SSVEP stimuli presentation.

To directly access the OpenGL API, to configure the environment, to open new windows or to draw objects, could be quite verbose from the source code point of view, thus various libraries, named toolkits, exist in order to provide handy functions leading to a little abstraction over the OpenGL API. The most famous and used toolkit is named OpenGL Utility Toolkit (GLUT) [86] and in particular its Open Source clone *Freeglut*. Anyhow, GLUT is a quite old project, no longer main-

tained and the Freeglut is mainly aimed at providing a stable clone of GLUT without newer or better features.

Other toolkits exist and for my implementation I chose to use one of the newest named OpenGL Framework (GLFW) ¹⁰ which in a similar application has been claimed to outperform GLUT [15], in particular for the timing functions which are able to exploit the best available timers on different operating systems [59].

The choice of all the used software tools, was aimed also to provide a flexible system able to run on different Operating Systems, with different hardware.

DOUBLE AND QUAD BUFFERING

Being OpenGL designed to draw not only static CG scenes, but also dynamic animations, it provides the capability to the programmer to draw every single frame which has to be shown on the screen. In OpenGL applications a *main loop* is commonly called infinitely in order to redraw the scene at every iteration to provide every time a new frame.

The scene redraw could be potentially very expensive from a computational point of view and thus it may take an amount of time which can not be considered infinitesimal. If only one buffer containing the scene pixels values would exist, it may happen that a scene could be displayed on the screen while it is still being drawn and thus scene contents may change while it is being displayed. To avoid this possibility, modern graphics cards implement a double-buffering system which can be managed through the OpenGL library.

In the OpenGL taxonomy two buffers named *front buffer* and *back buffer* are available and both of them can contain a scene, or in other words a frame and thus the color value of each pixel to be displayed. When redrawing a scene only the *back buffer* is used, while the *front buffer* is displayed on the screen. As soon as the screen, due to the refresh, needs a new frame, the *front buffer* and *back buffer* are swapped, in order to display the previous *back buffer* containing the new frame renaming it as the new *front buffer* and conversely letting the previous *front buffer* became the new *back buffer*. The new *back buffer* can therefore be cleaned and a new frame can be redrawn inside it.

¹⁰<http://www.glfw.org/>

The same approach is used also in what is called the *quad-buffering* technique, used to drive stereoscopic displays, where two different scenes have to be displayed, one for the left eye and one for the right one. In this case four buffers exist (i.e. *front left buffer*, *front right buffer*, *back left buffer* and *back right buffer*) and again at each screen refresh front buffers are swapped with back buffers. The reason for this additional piece of information will be clear in Sec. 6.2.

BUFFERS SWAPPING

Since, as already mentioned, the scene redraw may be computationally expensive, it would make no sense to iterate on the *main loop*, swapping the buffers, as fast as possible, because the screen refresh rate actually limits the number of displayable frames per second. Consequently, modern graphic cards and the OpenGL library, permit to access to a synchronization mechanism with the screen refresh, in order to redraw the scene only at the needed rate.

In particular, the scene redraw is triggered as soon as the buffers are swapped, since after the buffer swap a new *back buffer* is available to be erased and redrawn. The aforementioned synchronization mechanism is indeed implemented in the buffer swapping system in order to swap the buffers as soon as a screen refresh is requested and thus a new frame has to be displayed.

Therefore, the buffer swapping mechanism has to be perfectly synchronizable with the screen refresh and consequently it provides exactly the synchronization needed for the presentation of SSVEP stimuli. In particular, being the graphics card to control the screen refresh by hardware, this kind of synchronization is the best possible achievable by software, since OpenGL is directly accessing the graphics card driver. The same mechanism has indeed been successfully used for the SSVEP stimuli presentation in two recent works [15, 147].

In particular, using the GLFW toolkit, this synchronization mechanism can be exploited as shown in List. 5.1, where at line 2 the synchronization is enabled requesting the buffers swapping to happen every time the screen is refreshed with the parameter of `void glfwSwapInterval (int interval)`. If `interval` is zero, the swap will take place immediately when `void glfwSwapBuffers (void)` is called, without waiting for the screen refresh (also known as “vsync off” setting). Otherwise the

```

1 // Set how many frames to wait between each buffer swap
2 glfwSwapInterval(1)
3
4 // Main loop
5 do {
6
7     // Actually draw the scene in the GL_BACK buffer
8     drawScene();
9
10    // Swap front and back rendering buffers
11    glfwSwapBuffers();
12
13    // Count the frames
14    cicleNo++;
15
16
17 // Check for ESC key press or if the window was closed
18 } while( glfwGetKey( GLFW_KEY_ESC ) != GLFW_PRESS &&
19          glfwGetWindowParam( GLFW_OPENED ) );

```

Listing 5.1: A simplified portion of the *main loop* used to produce reliable flickering stimuli exploiting the OpenGL buffer swapping synchronization to the screen refresh.

`void glfwSwapBuffers(void)` function at line 11 became a blocking function, waiting at least `interval` screen refreshes to pass between each buffer swap (also known as “vsync on” setting). Using a swap interval of zero can be useful for benchmarking purposes, to measure the time needed to actually draw the scene.

In my implementation an `interval` set to 1 will always be used if not stated otherwise and thus the buffer swapping will occur at the same frequency as the screen refresh.

Being in this configuration `void glfwSwapBuffers(void)` blocking and being it synchronized with the screen refresh, at every iteration inside the *main loop* the exact frame number being drawn can be known simply counting how many time the `void glfwSwapBuffers(void)` is released, as can be seen in List. 5.1 at line 14.

TEXTURE CACHING

To augment the flexibility and versatility of the developed software, only simple objects are directly drawn using OpenGL (as quads) in contrast to what is done

in [147], while the texture functions available in OpenGL are used to change the aspect of the flickering stimuli, applying over the objects any picture or drawing the user my desire.

In my implementation the texture functions are used to implement the actual flickering of the stimuli, changing the texture applied over the objects according to the need of a stimulus onset or a stimulus offset; this is in some sense similar to what is proposed in [15]. Anyhow in contrast to what is done in [15], in my implementation is not used a large texture containing precomputed onset/offset patterns indexed as different sub-textures, but only an onset and an offset texture are loaded. The flickering pattern is generated “on the fly” counting the frame numbers to decide if the current frame requires an onset or an offset given the flickering frequency for the particular object and the screen refresh.

This leads to more flexibility, since my implementation can be run on different machines with different refresh rates without being modified and recompiled; moreover also the onset texture can be changed without recompiling, letting to be very handy to try different colors, shapes and pictures as stimuli.

Using the texture caching system, the onset and offset images are loaded at the very beginning of the program using the (Simple OpenGL Image Library) SOIL¹¹ library and then kept in the texture memory of the GPU. The offset picture is typically a black image or anyway a solid color image of the same color as the used background (in my experiments have always been solid black), but it could be used to perform experiments also using pattern reversal stimuli [207].

At each frame, as shown in List. 5.2, in the function drawing the scene, for every flickering object, a decision is taken in order to bind the onset or the offset image to the object. In the actual implementation the decision is taken evaluating the frame number given by `cicleNo` modulus the flickering ratio contained in the `ratio` variable, which was previously obtained as the screen refresh frequency divided by the flickering frequency¹².

Actually, to avoid variable overflows, the `cicleNo` variable is zeroed whenever

¹¹<http://lonesock.net/soil.html>

¹²Using this method only integer flickering frequencies could be obtained; this was not a limitation for my experiments, but if it is the case, different methods could be implemented in this part of the code to be able to use also all the possible lower frequencies.

```

1  if (cicleNo % ratio == 0) {
2
3      // Select OnSet Texture
4      glBindTexture(GL_TEXTURE_2D, texture[o]);
5      onsets++;
6
7  } else {
8
9      // Select OffSet Texture
10     glBindTexture(GL_TEXTURE_2D, texture[1]);
11
12 }

```

Listing 5.2: A portion of the function actually drawing the scene, reported to highlight the mechanism to bind to the on-set or off-set texture. In this simplified example the onset texture is showed for one frame only independently from the `ratio` variable, while in the actual implementation a flickering duty cycle of 50% (or as close as possible) is pursued.

it reaches a value corresponding to a common multiple of all the flickering ratios used in the scene.

OPENVIBE INTEGRATION

Apart from the stimuli presentation, the software I implemented has been integrated in the OpenVibe environment enabling it to send synchronization messages to be embedded within the EEG recordings.

In SSVEP BCIs, in particular when using self-paced BCIs, in general is not needed any synchronization between the stimuli presentation software and the EEG recording system. Despite of this, when analyzing data off-line, in particular when different stimulation methods have to be compared, to know exactly when in the EEG recording a stimulus was presented, could be extremely helpful. Moreover, if different stimuli have to be compared, it could be useful also to use different tags highlighting which stimulus was shown in each time window to the user. Furthermore, a precise software tagging could be useful also to perform time averages over multiple trials, as used in ERP experiments, to visualize the SSVEP waveforms.

Even more interestingly from the BCI applications point of view, a very precise software tagging could be exploited to implement *phase tagged* SSVEP based BCIs,

as introduced in Sec. 4.3.3. For example, in [209] where a LED stimulator was used to implement a *phase tagged* SSVEP based BCI using a single frequency, an hardware tagging system exploiting a photodiode had to be adopted. My software implementation, using software tagging, could avoid the use of external hardware such as the LED stimulator and the photodiode as well.

The OpenVibe *Designer*, as already mentioned, permits to use various boxes implementing different protocols to exchange information with other applications. Despite of this, these protocols were meant to send/receive control signals, but not precise synchronization triggers. Consequently they are not well suited to receive software triggers to be later embedded in the acquired data stream. Initially I implemented custom TCP and UDP servers to be added as OpenVibe tool boxes able to receive and send information to/from the *Designer*, but their latency limitations soon revealed these protocols to be not suitable for a very precise software tagging system.

Since it was in general a very interesting feature to have a faster protocol to exchange information with an external application, the OpenVibe developers added recently in the *Acquisition Server* the capability to receive triggering signals using the multi-platform Boost Inter-process communication library¹³. The only limitation given by this approach is the fact that the software sending the triggering signal has to be run on the same machine as the *Acquisition Server*; anyhow this is not a major limitation, since the *Designer* can anyway run on a different machine than the *Acquisition Server*.

Therefore I exploited this new capability to be able to send triggering signals, to be saved along with the EEG recording, from the developed stimuli presentation software, directly to the *Acquisition Server*.

In particular, to obtain the best precision possible, the trigger is sent as soon as the buffer are swapped, as shown in List. 5.3 at line 36. In the code fragment reported in List. 5.3 a software trigger seems to be sent every time the buffers are swapped, but actually in the `void sendStim(int stim)` function a control system is implemented in order to send a `OVTk_StimulationId_VisualSteadyStateStimulationStart` trigger only at the time the first onset texture is showed. Of course different behaviors could be implemented for different applications or debugging purposes; e.g.

¹³<http://www.boost.org/>

a trigger message could be sent at every buffer swap communicating if the current frame contains an onset or an offset for a particular object.

PERFORMANCE EVALUATION

To evaluate the performance of the stimuli presentation software the timing functions provided by the GLFW toolkit have been used, as have been done in [15] to accomplish the same task.

The test has been performed using the ASUS VG278 monitor described in Sec. 5.2.2 attached to a graphics workstation (DELL Precision T5600) equipped with an NVIDIA Quadro 4000 graphics card. The monitor was used as a regular monoscopic monitor at its maximum reachable refresh rate of 120 Hz.

The performed test was aimed to measure the precision and accuracy of the refresh rate and of the stimuli onsets, focusing in particular on their possible jitter. To perform the test the `double glfwGetTime(void)` function was used to access the system clock, which is claimed to reach resolutions in the order of 1 ns on modern PCs [59].

The time was measured after each buffers swap, while presenting flickering stimuli at different frequencies and was measured also every time there was an onset for each of the flickering stimuli. The results as mean and standard deviation are presented in Tab. 5.4.1 after being computed over 200 samples.

Refresh Rate		10 Hz OnSets		15 Hz OnSets		20 Hz OnSets	
Mean	Std	Mean	Std	Mean	Std	Mean	Std
8.335	0.0053	100.011	0.0022	66.674	0.0025	50.006	0.0020
119.98 Hz		9.999 Hz		14.998 Hz		19.997 Hz	

Table 5.4.1: In the first two columns are reported the mean time in milliseconds and its standard deviation between two buffer swap. Same values are reported in the subsequent columns regarding the time between two stimulus onsets for different stimulation frequencies. In the last row, the mean times are converted to the respective frequencies in Hertz.

As can be noticed, the timing is quite accurate and thus the performance level

```

1  #include "openvibeStimulationConnection.hpp"
2  #define OVTK_StimulationId_ExperimentStart    0x00008001
3  #define OVTK_StimulationId_ExperimentStop     0x00008002
4  #define OVTK_StimulationId_VisualSteadyState
5          StimulationStart 0x00008010
6  #define OVTK_StimulationId_Label_00          0x00008100
7  #define OVTK_StimulationId_Label_01          0x00008101
8
9  OpenViBE::StimulationConnection* osc;
10
11 void sendStim(stim) {
12
13     ...
14
15     osc->sendStimulation(OVTK_StimulationId_VisualSteadyState
16                          StimulationStart);
17
18     osc->sendStimulation(OVTK_StimulationId_Label_00+stim);
19 }
20
21
22 int main(int argc, char *argv[]){
23
24     ...
25
26     osc = new OpenViBE::StimulationConnection();
27
28     do {
29
30         ...
31
32         // Swap front and back rendering buffers
33         glfwSwapBuffers();
34
35         // Send software trigger
36         sendStim(stim);
37
38         ...
39
40     } while (...)
41
42 }

```

Listing 5.3: Code fragment of the developed software highlighting the software triggering mechanism sending a trigger signaling the start of a flickering stimulus and a label associated to the kind of stimulus.

seems to be adequate to provide reliable flickering stimuli to elicit precise SSVEP responses.

5.5 A COMPLETE SSVEP BASED BCI SYSTEM

The custom developed SSVEP eliciting stimuli presentation software described in Sec. 5.4.2, in conjunction with the custom feature extraction box described in Sec. 5.3.4 and the OpenVibe framework described in Sec. 5.3, permitted to implement a beyond state-of-the-art complete SSVEP based BCI system meant for the integration of flickering stimuli in generic VEs.

In Fig. 5.5.1 is shown a photo taken in the Virtual Theater, presented in Sec. 5.2.1, displaying a possible training phase for the custom implemented self-paced SSVEP based BCI system.

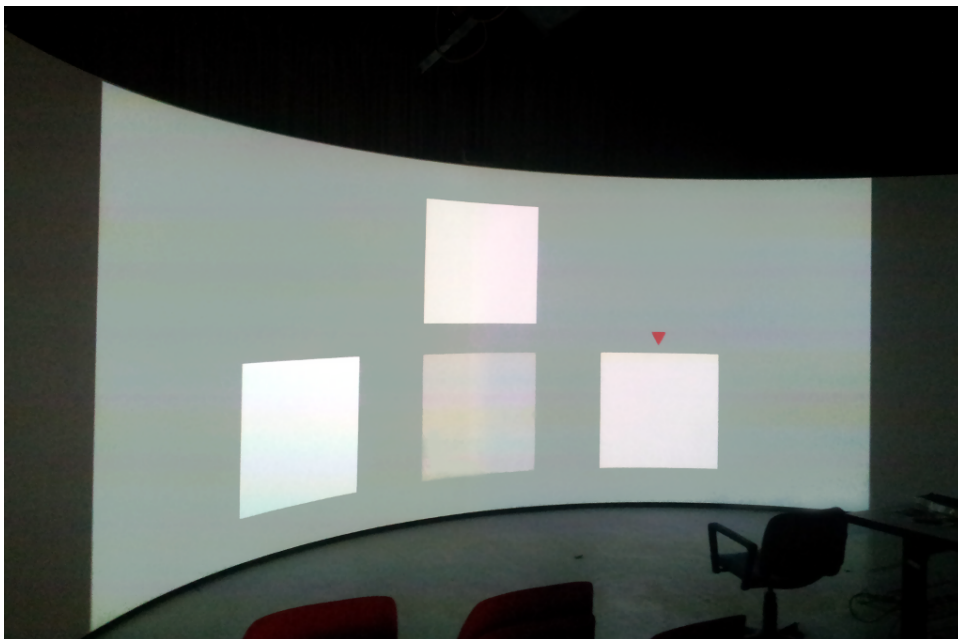


Figure 5.5.1: A possible stimuli presentation for the training phase of the custom implemented SSVEP based BCI.

In this generic BCI example, used for system testing purposes, 3 white flickering targets, plus a non-flickering one are shown to the user over a gray background. Being the stimuli presentation software the OpenGL code discussed in Sec. 5.4.2,

the aspect of the presented scene can be easily customized and integrated in pre-existing VEs.

During the training phase a red triangle indicates to the user which target to attend while the EEG signal is recorded along with the timestamps corresponding to the gazed target. All the communications between the stimuli presentation software and OpenVibe were implemented using the software triggering method described in Sec. 5.3.1

Using the OpenVibe *Designer*, two different scenarios have been created, exploiting the custom implemented features extraction box.

The first scenario, shown in Fig. 5.5.2, is used to train the Linear Discriminant Analysis (LDA) classifiers to separate the features corresponding to the different attended stimuli. At the moment only two class classifiers are available in the OpenVibe framework, thus multiple classifiers have to be used.

In particular, for the example shown in Fig. 5.5.2, used in conjunction with the stimuli presentation shown in Fig. 5.5.1, four classifiers are trained; three of them are trained in order to separate epochs containing a SSVEP response to one of the three flickering frequencies, plus another classifier trained to recognize an EEG epoch containing no responses to any of the used frequencies.

Scope of the scenario shown in Fig. 5.5.2 is to separate the different epochs accordingly to the attended target; for each epoch the feature vector is computed using the custom implemented Minimum Energy Combination box and then each classifier is fed with the feature vectors corresponding to its two classes.

The second scenario, shown in Fig. 5.5.3, to be run on-line while using the BCI, exploiting the previously trained classifiers determines the attended target and provide a feedback to the user. In this particular example the feedback is provided as a synthetic voice telling the user the position of the target detected by the system (“left”, “center” or “right”). Being the implemented BCI system self-paced, the audio feedback to the user is provided only when the system detect a significant SSVEP response and not on a time periodic basis.

This scenario just computes the T index for each signal epoch, where both the epoch length and the overlap factor are configurable. The T index is computed

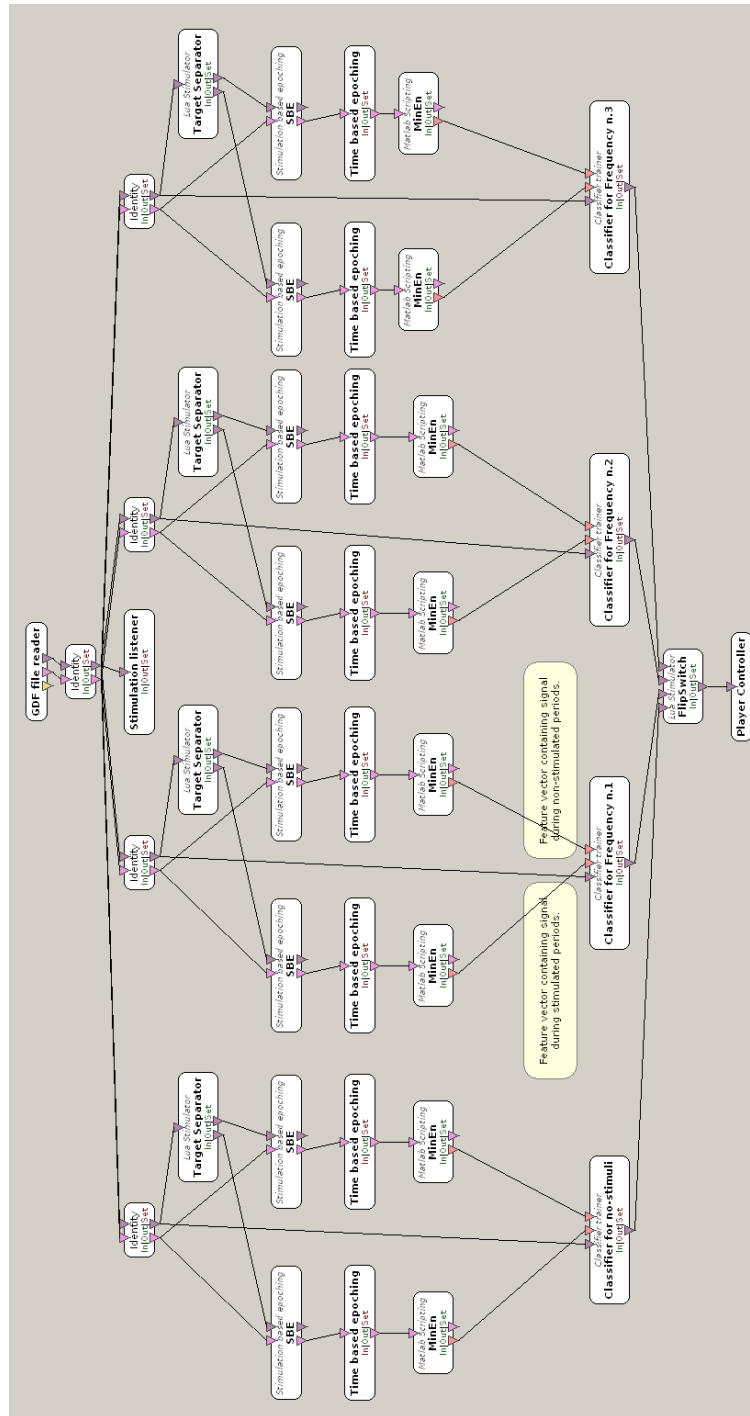


Figure 5.5.2: Training scenario of the custom implemented SSVEP based BCI exploiting the Minimum Energy Combination method.

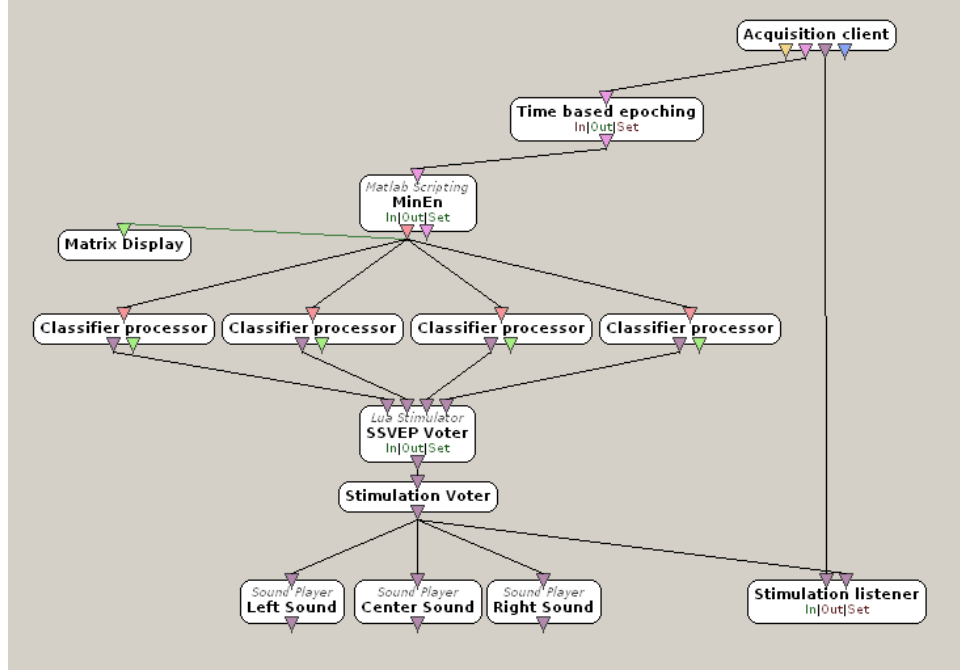


Figure 5.5.3: On-line scenario of the custom implemented SSVEP based BCI exploiting the Minimum Energy Combination method.

for all the used flickering frequencies, providing to the classifiers a feature vector having a dimension corresponding to the flickering targets, e.g. $\langle T_{f_1}, T_{f_2}, T_{f_3} \rangle$.

Then, each classifier provides the computed class for each feature vector, telling if the feature vector is considered to contain a response to the flickering frequency associated to the classifier. In this particular implementation one of the classifiers is used to identify features vector containing no-response; this is redundant since it correspond to the event when none of the three classifiers associated to the flickering frequencies detect a response, but having a fourth classifier trained to this purpose demonstrated to increase the avoiding of false detections.

The results of all the classifiers are then evaluated and merged to obtain a single result (in the *SSVEP Voter* box shown in Fig. 5.5.3): if only one classifier answers with a detection, that will be the global result, otherwise if more than one classifier detect a response, the global answer became a no-response detection.

Eventually, in the *Stimulation Voter* box, shown in Fig. 5.5.3, a sequence of results corresponding to subsequent epochs are evaluated thanks to a voting mechanism where if the absolute majority of the results correspond to a particular flickering

frequency a target detection is triggered. Otherwise, if an absolute majority can not be found no detection is triggered.

5.6 CONCLUSION

In this Chapter the available hardware tools have been described and moreover the performed software developments have been discussed.

All of the presented software developments aimed to the implementation of a beyond state-of-the-art SSVEP based BCI system for the integration of flickering targets in a generic VE.

The engineering efforts have been considerable because of the highly heterogeneity of problems to be faced, requesting different solutions spanning in different research areas. Several languages and libraries have to be used, as well as different software frameworks, in order to solve issues concerning visual stimuli synchronization, software triggering, data acquisition, data processing, feature extraction and machine learning.

The obtained system can be considered beyond the state-of-the-art since it improves the already available SSVEP based BCI scenarios dispatched within the OpenVibe framework.

In particular, a first improvement is given by the implementation of the custom stimuli presentation software described in Sec. 5.4.2, which provides reliable flickering stimuli, high precision software triggering and moreover an high flexibility, being able also to provide particular flickering stimuli exploiting stereoscopic displays, as will be detailed in Sec. 6.2.

A second improvement is given by the implementation of the Minimum Energy Combination feature extraction method, which allowed to reduce the training time and to improve the system performances. The training time is reduced since spatial filters training is not needed when using this method and thus only two scenarios are needed; one to train the classifiers and one for the actual BCI. On the other side, also the BCI performance is improved, as discussed in Sec. 4.3, since this method is able to separate the actual SSVEP response from the uncorrelated brain activity occurring in the same frequency bands of the stimuli.

6

Performed Experiments

In this chapter will be described the performed experiments, which will be divided in three sections, tackling different challenges regarding the use of SSVEP based BCIs in the context of VR environments and Computer Games.

The first Section will deal with the necessity of commodity, low-cost, portable and usable devices, able to detect the SSVEP response for practical SSVEP based BCI applications.

The second Section will deal with the possibility of using stereoscopic display devices do provide ordinary SSVEP eliciting stimuli in the context of stereoscopic environments, but also with an experiment highlighting the capabilities of these devices exploitable to elicit particular SSVEP responses.

Eventually, in the last section are proposed different experiments investigating novel approaches in order to exploit the SSVEP response to implement prospective *Passive* BCIs.

6.1 SSVEP BCI USING THE MINDSET

To move towards dry electrodes system is one of the most challenging research directions for practical BCIs in the last years [34, 116, 117].

Gel based electrode systems are the state-of-the-art in terms of signals acquisition quality, but the electrodes montage often requires a trained technician. Just the fact that the user would need another person to mount the EEG headset on him/her is a strong limitation for end-user applications.

Custom devices are being investigated [204] in order to provide easier to use devices by means of dry electrodes or salted-water based electrodes and also industrial companies are working towards this direction (as Emotiv, NeuroSky, Biosemi, g.Tec, etc.), despite of this, no commercial devices explicitly aimed for SSVEP based BCI are yet available implementing these technologies.

Professional commercial general purpose EEG headsets implementing dry electrodes systems are available (e.g. the g.Tec g.SAHARA system ¹), but from their cost and their complexity (e.g. the need to chose electrodes positions, connect the electrodes to the amplifiers, etc.) is clear that they are meant to be used by researchers and not end-users yet.

The MindSet device presented in Sec. 5.1.1 is clearly not designed to detect SSVEP responses, but in the view of “more practical BCIs”, it would be very interesting if the MindSet could be used also for SSVEP based BCIs, thanks to its extremely low cost and for its ease to wear. In particular, in contrast to the *Active* BCI its manufacturer had in mind, it would be very interesting to be able to use it for *Reactive* BCIs based on the SSVEP modality, since it would avoid the need of subject training. Applications where a generic user would just need to wear the device in order to be able to use it, using a *Reactive* SSVEP based BCI approach, would be possible.

Other commodity hardware devices have been recently successfully used to implement SSVEP based BCI, as for example the Emotiv EPOC ² device [106], which is a salted-water based 14-electrodes system.

¹<http://www.gtec.at/Products/Electrodes-and-Sensors/g.SAHARA-Specs-Features>

²<http://www.emotiv.com/epoc/>

Anyhow for the MindSet the challenge is quite harder: having a single electrode it will not allow to apply spatial filters algorithms, as the ones exposed in Sec. 4.3.1; having a dry electrode it will provide a noisier signal [116] and moreover, as mentioned in Sec. 5.1.1, wearing it as indicated by its manufacturer, the electrode would be positioned roughly at Fp1, that is on the forehead and thus very far from the visual cortex where SSVEP responses are more intense.

6.1.1 MATERIAL AND METHODS

For all the tests performed in this section have been used the custom developed stimuli presentation software presented in Sec. 5.4.2. For the EEG signal acquisition, from the hardware point of view has been used the MindSet device, presented in Sec. 5.1.1, while from the software point of view the OpenVibe framework, presented in Sec. 5.3.

To overcome the wrong positioning of the MindSet's electrode for SSVEP response detection, I have conducted the experiments using the MindSet on the subjects' head, swapping the left and right headphones, letting the single electrode to be positioned backwards. In this manner, as shown in Fig. 6.1.1, the electrode is roughly positioned near P2 (according to the extended 10-20 system), which is a much more suitable location to detect the SSVEP response, although not the optimal one [188].

Being the MindSet electrode meant to be positioned on the forehead, its shape is not appropriate to have a connection to the scalp where hair is present. To overcome this problem a droplet of conductive electrode gel was used to improve the contact. Although consequences in terms of impedance could not be assessed due to the proprietary hardware (electrode-skin impedance could not be measured and the amplifier input impedance is unknown), experimental results, as exposed in the next sections, confirm that this procedure improve the acquisition signal quality. In order to have a similar impedance also on ground and reference electrodes, a very small amount of gel was positioned also on them.

Despite of the use of a little amount of gel, wearing the MindSet remains much easier than wearing other gel based EEG devices (e.g. the g.MOBILab+) and it can be easily done by the subject with no need of external help. Moreover, the small



Figure 6.1.1: Figure depicting an user wearing the NeuroSky MindSet device “reversed”, with its electrode backward facing, positioned roughly at P2 location (according to the extended 10-20 system), over the parietal lobe of the cerebral cortex.

amount of gel to be used, do not force the subject to have a shower right after the use of the device. Furthermore, the user can use the automatic impedance checker, implemented in the device, in order to assess the quality of the electrodes connection which is reported in real-time in the *OpenVibe Acquisition Server*. Montage of the device is in the range of about 0.5 min to 3 min, according to the hair volume of the users.

6.1.2 RESULTS OF THE PRELIMINARY EXPERIMENT

As can be seen in Fig. 6.1.2 and in particular in Fig. 6.1.3, wearing the MindSet with backward facing electrode, in a preliminary experiment, lead to a clear recording of a SSVEP response. The subject was looking at a 15 Hz flickering white patch displayed on a regular 60 Hz LCD screen (described in Sec. 5.2.3) and the PSD was obtained applying the Wiener-Khinchin theorem and thus taking the Fourier Transform of the 60 s long signal auto-correlation ³.

³In this first trial the duty cycle of the stimulus was not tuned and was 25% (one onset frame out of four for every period).

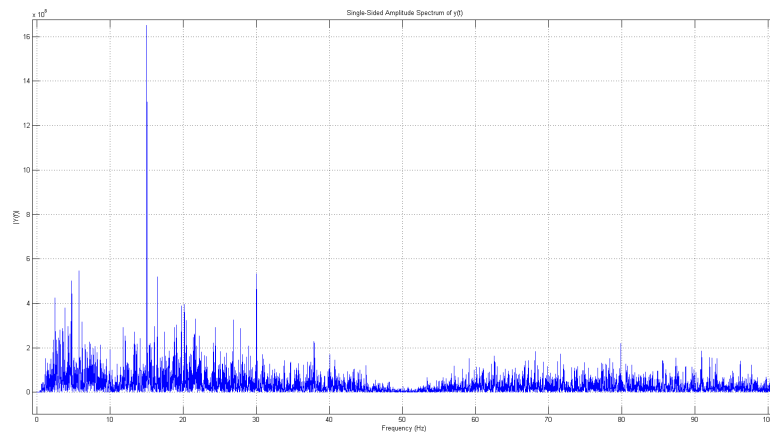


Figure 6.1.2: The PSD spectrum of the EEG signal acquired by the single dry electrode MindSet device positioned on my scalp while attending a 15 Hz flickering pattern on a regular 60 Hz screen for 30 s.

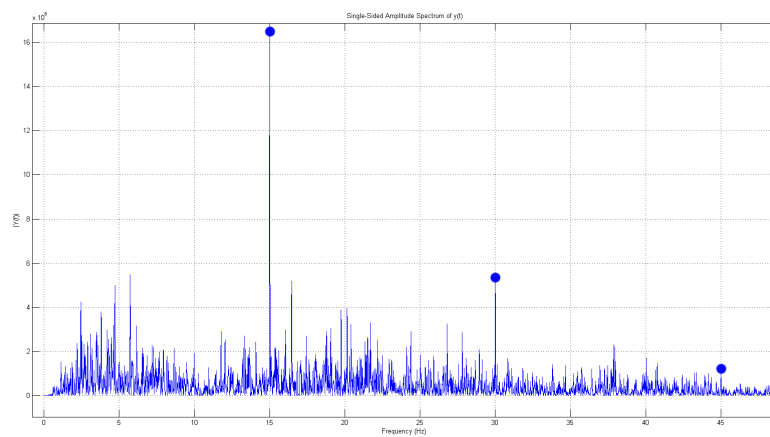


Figure 6.1.3: The same spectrum as in Fig. 6.1.2, but enlarged between 0 Hz to 50 Hz and with added blue points highlighting the height of the peaks at the fundamental frequency and its first harmonics.

6.1.3 RESULTS OF THE OFF-LINE CLASSIFICATION

Knowing that the SSVEP response could be recorded with the MindSet, further experiments have been conducted to identify the shortest signal length able to lead to a classification accuracy high enough for BCI applications.

In a similar fashion as done in [53], 30 s of data was recorded for every trial for two different stimulation frequencies, chosen as 12 Hz and 15 Hz. Four trials have been performed for each subject in order to have a total of 60 s recording for each frequency.

The duty cycle of the stimulus has been tuned for best performances according to the results discussed in Sec. 4.2.2 and was 50% for the 15 Hz stimulation frequency and 40% for the 12 Hz one (since in this case a 50% duty cycle is not possible on a 60 Hz display due to the integer odd number of frames).

Off-line analysis has later been performed using the SSVEP detection method proposed in [53] and discussed in Sec. 4.3.2, in order to estimate the ratio between the SSVEP response and the uncorrelated brain activity, in correspondence of each of the stimulation frequencies, for every 1 s and 2 s of non-overlapping signal windows, computing the T index ⁴.

As shown in Fig. 6.1.4, where data from a preliminary experiment is reported, using two second signal windows a linear classification between the epochs acquired under two different stimulation frequencies seems to be feasible, despite of the quite short signal windows, considering the used acquisition device.

Using the same approach, the T index has been computed on the same dataset also for one second time windows and the results are shown in Fig. 6.1.5. As can be seen in this case a linear classification would produce a lower accuracy, but it seems to be possible anyway.

Multiple experiments were performed on a set of subjects to assess the reachable classification accuracy on a larger population. The classification has been performed with a software tool, provided by Lorenzo Rosasco ⁵ during a graduate

⁴The Minimum Energy Combination method has not been applied before the SSVEP response detection, since having only one electrode signal there are no possible signals combinations apart from the trivial one.

⁵<http://web.mit.edu/lrosasco/www/>

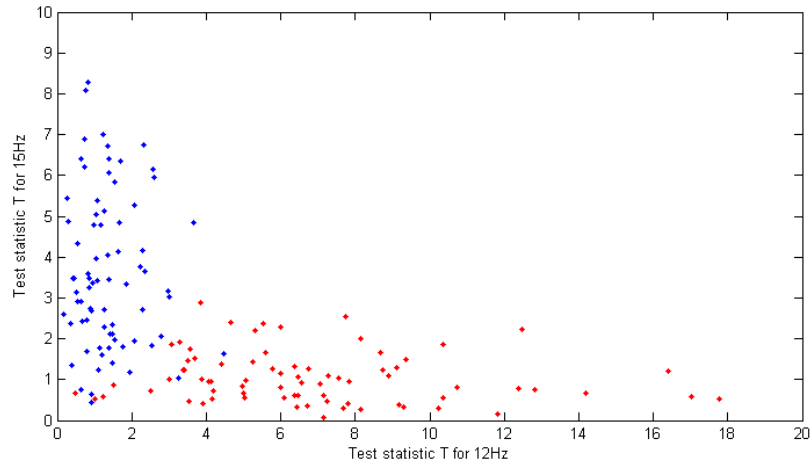


Figure 6.1.4: T index computed for the two frequencies for every two seconds non-overlapping window of EEG signal. Blu points refer to epochs with a 15 Hz stimulation, while red points refer to epochs with a 12 Hz stimulation. Plotted data correspond to 10 trials acquired from one subject, for a total of 150 epochs (75 for each frequency).

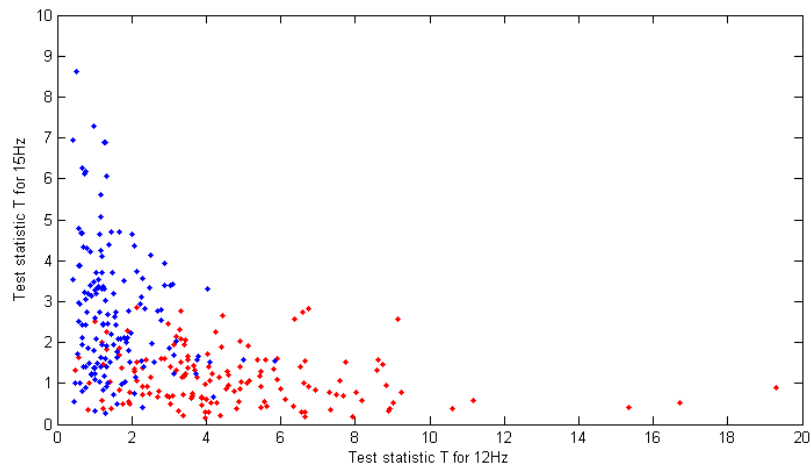


Figure 6.1.5: T index computed for the two frequencies for every one second non-overlapping window of EEG signal. Blu points refer to epochs with a 15 Hz stimulation, while red points are epochs with a 12 Hz stimulation. Plotted data correspond to 10 trials acquired from one subject, for a total of 300 epochs (150 for each frequency).

student class about machine learning. The used software implement various spectral regularization methods for supervised learning and is able to show on a graph the computed classification applied on training and test data for bidimensional datasets. For this work, due to the nature of the data to be classified, has been used a linear least squares binary classification. The main window of the software showing a loaded dataset used for the classifier training is showed in Fig. 6.1.6.

The classification could be easily computed with other simpler software tools, but this has been chosen in order to be able to plot and visually inspect the datasets and their classification.

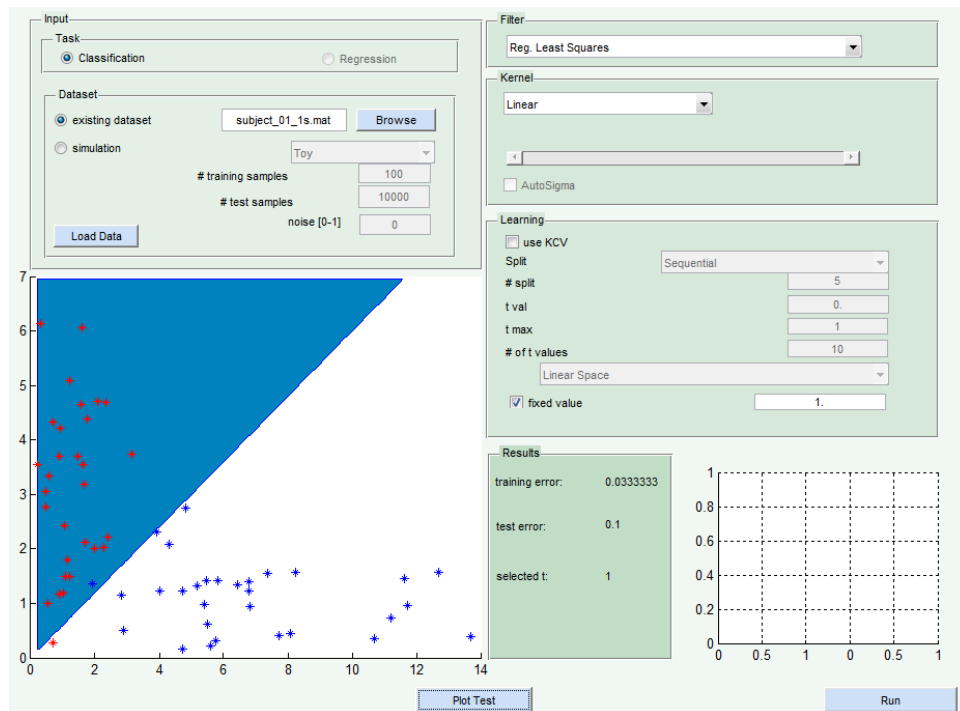


Figure 6.1.6: Graphical User Interface of the software used to train the classifier on the train dataset, to classify the test dataset and to compute the train and classification accuracies. In the bottom-left part of the GUI can be seen the training dataset for the Subject 1 (1s windows) and the computed line separating the two class which will be used to classify the test dataset.

The classifier training was performed using one trial for each of the two frequencies, while the remaining two trials where used as test data; results for the different

	One second windows		Two seconds windows	
	Train Error	Class. Accuracy	Train Error	Class. Accuracy
Subject 1	0.03	90%	0.00	93%
Subject 2	0.15	83%	0.03	90%
Subject 3	0.43	74%	0.40	87%
Subject 4	0.33	70%	0.33	80%
Subject 5	0.46	69%	0.30	83%
Subject 6	0.40	50%	0.33	48%

Table 6.1.1: Classifier train error and classification accuracy computed using a linear Least Squares classification. Results are reported for 1 s and 2 s signal window length, computed on the same dataset. The classification accuracy gives the predictability of which stimulation frequency the user was attending to (between 12 Hz and 15 Hz), given one signal window.

subjects are reported in Tab. 6.1.1.

According to the reported results, for 5 subjects out of 6 the SSVEP response could be detected with a reasonable accuracy. As expected, using 2 seconds epochs lead to better results for all the subject apart for the 6th one. For the 6th subject, a manual inspection of the data points revealed that the two point clouds, relative to the two stimulation frequencies, are not separable for all the acquired trials. The reason may be a SSVEP BCI illiteracy of the subject, a very low attention payed to the flickering target or a particularly inefficient electrode location for the particular subject.

Concerning the classification accuracy, it is worth to mention that it has been computed using non-overlapping windows, but, when implementing SSVEP based BCIs, is a common practice to compute the SSVEP response index (e.g. the narrow-band power or the T index as in my case) for sliding windows and then to evaluate the computed value for several subsequent windows. This leads to a smoother output removing the effect of “false-positive detections” which may be computed in a single signal window. I did not used this approach to compute the values in Tab. 6.1.1, since it would have been not a mathematically correct way to evaluate the classification accuracy, in the sense that multiple points would have been computed from the same parts of acquired signal.

Despite of this, in a real application, using sliding windows, the accuracy is ex-

pected to be the same or higher.

6.1.4 RESULTS OF THE ON-LINE ACTUAL BCI

Given the promising results reported in Tab. 6.1.1, the MindSet has been used to perform a new experiment to test its performance in an actual BCI implementation.

Five subject wearing the MindSet used the custom self-paced three-targets SSVEP based BCI system described in Sec. 5.5.

Three flickering targets at 10 Hz, 12 Hz and 15 Hz, plus a non-flickering one, were presented in the Virtual Theater as shown in Fig. 5.5.1 for classifiers training. The training phase was 96 s long, 6 s of EEG data was acquired for each target (three flickering ones, plus the non-flickering one), for 4 repetitions, giving a total of 24 s data for each target. Two second overlapping epochs were used, computed every 0.125 s.

As detailed in Sec. 5.5, four LDA classifier were trained, where one classifier separates the feature vectors containing no SSVEP responses from the ones containing it, while the other three classifiers separate the feature vectors containing a response to a specific flickering frequency from the ones not containing it.

The training results of the four LDA classifiers, using 6 partitions k-fold cross-validation, for the 5 subjects are reported in Tab. 6.1.2.

	10 Hz	12 Hz	15 Hz	no-stim
Subject 7	91.67%	100.0%	88.89%	80.56%
Subject 8	86.11%	94.44%	86.11%	83.33%
Subject 9	83.33%	94.44%	83.33%	61.11%
Subject 10	75.00%	94.44%	66.67%	69.44%
Subject 11	66.67%	88.89%	77.78%	58.33%

Table 6.1.2: Training results of the four LDA classifiers using a k-fold cross-validation with 6 partitions. The first three columns represent the detection accuracy for the respective flickering frequencies targets, while the fourth column represent the detection accuracy of the absence of a SSVEP response.

The on-line scenario was configured in order to compute a feature vector every

0.125 s using 2 s epochs and to classify it in parallel with all the classifiers. All of the classification results were then merged as described in Sec. 5.5 and subsequently a majority voting mechanism was applied as soon as at least 16 results (corresponding to 2 s of EEG data) were available.

The same scene, shown in Fig. 5.5.1, used for training, was later presented to all the subjects to test the on-line actual SSVEP based BCI.

The use of a fourth classifier to detect the absence of a SSVEP response lead to a very low error rate in the target detection, but on the other side, for some subjects it increased the time needed to have a response from the system. In particular, for subject 10 and subject 11 the system used to be silent for more than two seconds after a gaze shift between different targets, before answering with the correct target detection.

On the other side, the first three subjects were able to use the system at its maximum speed (one target detection every 2 s), with no errors, apart from the errors due to the 2 s latency when changing the gazed target.

6.1.5 CONCLUSION AND FUTURE WORKS

With the previously described experiments it has been demonstrated that using a popular single electrode consumer-grade EEG acquisition device is possible to detect a SSVEP response.

Moreover, despite of the not optimal electrode position and its physical shape, it has been demonstrated that, using a state-of-the-art signal processing technique, the signal window length needed to accurately detect the SSVEP response could be short enough for BCI applications.

Furthermore, a full SSVEP based BCI has been implemented and tested using the MindSet device, showing that despite of the low cost of the device, thanks to the precision of the stimuli presentation software, to the performances of the feature extraction algorithm and to the flexibility of the OpenVibe framework, it can be used for SSVEP based BCI applications.

The reported results highlight the feasibility to implement a SSVEP based BCI using the MindSet device and the presented signal processing method. This is interesting due to the wide diffusion and affordable cost of this device, but more

importantly for its ease of use.

This result is extremely interesting for applications aimed to the interaction of healthy end-users with Computer Games and VE, but also for patients needing BCIs for every-day use, where complex and expensive EEG acquisition devices are not handy and a lower accuracy can be interestingly counterbalanced by a more convenient solution.

Even more interestingly, this work highlights the possibility to design new affordable single electrode devices, specifically for SSVEP based BCI applications, adopting a more suitable electrode position and a specific electrode shape to let it be positioned where hair is present, without the need of conductive gel and without affecting the device cost.

6.2 SSVEP ELICITATION BY MEANS OF STEREOSCOPIC DISPLAYS

Stereoscopic visualization in cinematography and Virtual Reality (VR) creates an illusion of depth by means of two bidimensional images corresponding to different views of a scene.

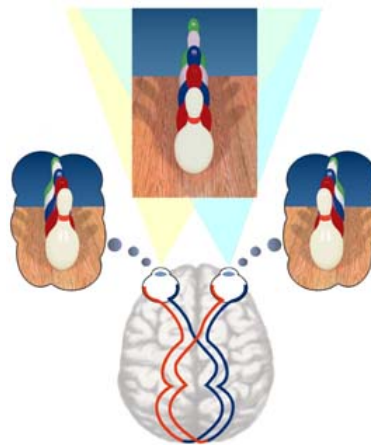


Figure 6.2.1: Simple representation of the depth perception given by the two different perspective views in the two eyes. Figure taken from <http://www.vision3d.com/>.

This illusion is based on making believe to the Human Visual System (HVS) that the two different images correspond to the two different perspective views

captured by the eyes, as shown in Fig. 6.2.1. The two images commonly contains different objects which, according to their depth, present different horizontal disparities between the two views.

The introduction of stereoscopy in the production of movies or of VR environments allows an observer to enhance the sense of immersivity and presence. In the context of virtual reality installations, stereoscopic devices are indeed widely used and recently they were introduced also in the end-user market as stereoscopic televisions and computer monitors.

The use of BCI systems in conjunction with VR environments demonstrated to produce various benefits, as described in Sec. 3.2.3 and VR environments are often presented to the users by means of stereoscopic displays. Despite of this, very few studies are available about the influence that stereoscopic displays may have on the SSVEP stimuli presentation and consequently on the SSVEP elicitation and detection.

An interesting question regarding these display devices, is whether or not they could be used to integrate SSVEP stimuli inside the presented stereoscopic VE, and, more interestingly, if their capability of showing two independent images to the user's eyes could be exploited to enhance SSVEP based BCIs performances.

Dichoptic stimulation ⁶, both for VEP and SSVEP elicitation, was used mainly in the field of vision research in the eighties and nineties, with the aim of studying how the depth perception given by binocular disparity is handled by the HVS [73].

SSVEP based BCIs were implemented in the past also using stereoscopic displays, thus it is known to be possible to use them to this aim, but few studies compared the same BCI both on monoscopic and stereoscopic displays concerning the effectiveness of the presented stimuli. Moreover, as far as I know, there are no research works trying to exploit them to provide dichoptic stimuli aimed to implement better SSVEP based BCIs.

In a very recent work [125] a SSVEP based BCI, implementing a navigation task, has been compared for a set of subjects, using a monoscopic and a stereoscopic VR environment. Some of the users performed better in the former condi-

⁶Dichoptic refers to viewing a separate and independent field by each eye. In dichoptic presentation, a stimulus A is presented to the left eye while a stimulus B is presented to the right eye.

tion, while others in the latter. Anyhow, interestingly the authors found a correlation between the user performances and their reported visual fatigue, highlighting the fact that the benefit of using a more immersive environment is counterbalanced by an increased visual fatigue for the group of most sensitive users.

An increased fatigue for the user can indeed have, as a consequence, a decrease in the attention paid to the stimuli which, as detailed in Chap. 4, is essential for a strong SSVEP elicitation, thus provoking a decrease in the overall BCI performance.

Unfortunately, in the mentioned research work is not clear if also the flickering stimuli were presented with a non-zero stereoscopic disparity or not.

In the first experiment I present in this section, my aim is indeed to study if a flickering stimulus presented with a non-zero stereoscopic disparity, independently from any other depth cue, elicits a stronger or a weaker SSVEP response in the subjects, than when presented with a zero stereoscopic disparity. This would be an interesting information to take into account when planning SSVEP based BCIs in the context of VR environments with tightly integrated visual stimuli as the one presented in [105].

On the other side, in the second experiment, my aim is to identify a possibility to exploit the capability of showing dichoptic stimuli, offered by stereoscopic displays, not only to generate the illusion of depth in the user, but also to overcome the limitation to the usable number of targets in a *frequency tagged* SSVEP based BCI.

From previous experiments in the field of vision research, is known that using a stereoscopic target which change at a constant frequency its depth ⁷, instead of its color or pattern, is possible to elicit VEP [73] and consequently probably SSVEP as well.

Despite of this, the amplitude of the evoked VEP was reported as being much smaller than using a color or pattern change, thus this seems not an efficient way to exploit stereoscopic displays for better SSVEP based BCIs. Moreover a change in stereoscopic disparity is commonly associated to muscular eye movements due to

⁷As a depth change is meant a change in the stereoscopic disparity between the stimuli presented to the two eyes.

eye vergence ⁸ which are known to produce EEG artifacts [73], but also to easily elicit fatigue in the users.

A more interesting approach, made possible by the flexibility of the implemented stimuli presentation software, would be to use different stimulation frequencies for the two different stereoscopic views of the same target. If such frequencies combinations would be detectable in the EEG signals, for example, three different targets could be presented using only two frequencies f_1 and f_2 where: target 1 flickers in both views at f_1 , target 2 flickers in both views at f_2 while target 3 flickers at f_1 in one view and at f_2 in the other.

The use of a dichoptic stimulation using two different flickering frequencies for the same target should in principle present in the EEG non-linear combinations of the used frequencies as discussed in Sec. 4.2.8 and presented also in [33] for multiple flickering frequencies/colors.

The use of this kind of stimulation was indeed discussed in a vision research work investigating the detection of VEP's non-linear components elicited by a dichoptic two frequency stimulation [162]. These kind of non-linear interactions were later studied also in [206] and furthermore in [171] using the bispectrum higher order spectral analysis.

Thanks to these studies, as recalled also in [73] ⁹, is known that showing to the left eye of a subject a flicker of frequency f_1 and to the right eye a flicker of frequency f_2 , non-linear processes produce harmonics of f_1 in left eye pathway and harmonics of f_2 in the right eye pathway. Moreover, non-linear processes, occurring after the monocular signals are combined, produce cross-modulation terms of the general form $nf_1 + mf_2$ for integral values of n and m .

In the second experiment presented in this section I therefore propose a novel approach to exploit stereoscopic displays to present this kind of dichoptic stimulation to overcome the limited number of targets presentable in a *frequency tagged* SSVEP based BCI. Such a paradigm could be helpful also to use a small set of frequencies, known to elicit strong SSVEP responses, to present a larger set of targets.

⁸Vergence is the simultaneous movement of both eyes in opposite directions to obtain or maintain single binocular vision. The eyes must rotate around a vertical axis so that the projection of the image is in the center of the retina in both eyes.

⁹In particular, Sec. 11.7 and Sec.13.1.8b

6.2.1 MATERIAL AND METHODS FIRST EXPERIMENT

Data acquisition was performed with the g.Tec g.MOBilab+ multipurpose version described in Sec. 5.1.2 and the OpenVibe framework was used to store EEG traces to file, within stimulus start/end triggers, for off-line analysis.

Electrodes was placed on 4 scalp locations over the visual cortex, POz, Oz, O1, O2, referenced at the left ear lobe and grounded at Fpz according to the extended 10-20 system.

Using the custom developed software presented in Sec. 5.4.2 and the stereoscopic monitor presented in in Sec 5.2.2, a single 15 Hz flickering square white patch over a black background was shown to each subject.

The stimulus was shown for 8 s with zero disparity between the two views and then followed by 5 s of no-stimulus, then the sequence was repeated with the same stimulus, but with a positive disparity between the two views and later again with a negative disparity. The whole trial was then repeated 4 times for each subject.

Stimulus size was kept in the order of 6° of visual angle adjusting the users distance from the monitor, while both positive and negative disparities where chosen to not induce eye strain, but anyhow to be relatively strong and thus in the order of 1° of visual angle.

The SSVEP response was later computed for all the EEG epochs where a stimulation was present using the Minimum Energy Combination method and the T index [53], described in Sec. 4.3.1.

6.2.2 RESULTS OF THE FIRST EXPERIMENT

The results of this experiment highlighted that a significant correlation between the stimulus stereoscopic disparity and the SSVEP response strength is not present.

Mean and standard deviation of the computed T index are reported in Tab. 6.2.1 where an inter-subject variability can be appreciated, but where the disparity between the two views of the stimulus seems not to influence the strength of the elicited SSVEP response.

It is worth to notice that the stimulus size was kept fixed and only the disparity between the two views was changed, thus most of the users did not perceived any

	Negative Diparity		Zero Disparity		Positive Disparity	
	Mean	Std	Mean	Std	Mean	Std
Subject 1	7.85	0.49	8.47	1.78	9.05	0.21
Subject 2	4.83	0.78	3.54	0.36	5.82	0.61
Subject 3	4.50	2.20	6.56	3.12	3.75	2.10
Subject 4	4.13	1.10	3.66	0.95	4.10	0.30

Table 6.2.1: SSVEP T index computed for 4 subjects attending the same flickering stimulus at 15 Hz presented on a stereoscopic display with different disparities.

difference between the different disparity stimuli, apart from the sensation of the eyes moving due to vergence. This is to highlight the fact that using other depth cues as objects size changes and perspective views, a conscious depth perception may influence the user attention and consequently the SSVEP response significantly.

Anyhow, aim of this experiment was to assess only the possible influences given by the dichoptic stimulus disparity. More complex setups would poses much higher challenges in order to separate the effects of the different phenomena involved; e.g. a stimulus size change would for sure influence the SSVEP response, but it would be an independent phenomena with respect to the stimulus disparity.

6.2.3 MATERIAL AND METHODS SECOND EXPERIMENT

In the second experiment the same hardware setup and electrode locations were adopted. The same stimuli characteristics were used as well, apart from the fact that in this case three patches where present on the screen, one on the left, one in the center and one on the right.

The left patch was flickering at 15 Hz, while the right one at 20 Hz for both of the stereoscopic views. On the other hand the patch in the center was presenting a dichoptic flicker showing to the left eye of the subject a patch flickering at 15 Hz and on the right eye a patch flickering at 20 Hz.

The subjects were asked to attend for 60 s each of the three patches as shown in Fig. 6.2.2 and the corresponding EEG signals were saved on a file and tagged with



Figure 6.2.2: One subject performing the experiment. Left and right patches are flickering respectively at 15 Hz and 20 Hz, while the center one is providing a dichoptic flicker showing to the left eye of the subject a patch flickering at 15 Hz and on the right eye a patch flickering at 20 Hz.

the corresponding patch identifier.

The same experiment has been later performed also using two closer flickering frequencies corresponding to 12 Hz and 15 Hz.

6.2.4 RESULTS OF THE SECOND EXPERIMENT

The PSD of the recorded files was computed as the Fourier Transform of the 60 s long signal auto-correlation, in order to highlight the contained frequency components.

As can be appreciated from Fig. 6.2.3, the PSD of the signal coming from one of the electrodes (Oz in this case) while the subject was attending the left and the right patches, as expected, reports a clear peak respectively at 15 Hz and 20 Hz.

On the other side, from Fig. 6.2.4 can be appreciated that the SSVEP response elicited by the dichoptic stimulus presented in the center of the screen, in its PSD contains different peaks which were not present before.

These peaks clearly denote a non-linear interaction between the SSVEP elicited

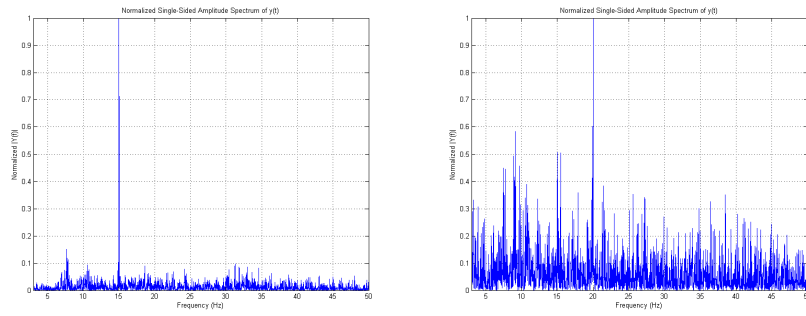


Figure 6.2.3: The normalized PSD of the EEG signal acquired by the Oz electrode while the subject was attending a flickering white patch. The attended patch was the left one (flickering at 15 Hz), on the left of the image, while it was the right one (flickering at 20 Hz), on the right of the image. The apparent different intensity of the background EEG is due to the normalization, since for this subject/electrode/stimulus combination the 20 Hz peak is weaker than the 15 Hz one.

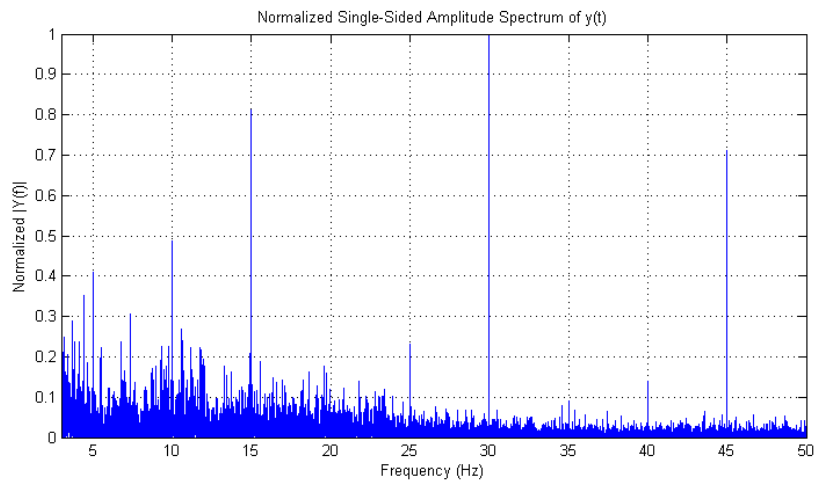


Figure 6.2.4: The normalized PSD of the EEG signal acquired by the Oz electrode while the subject was attending a dichoptic flickering white patch, with stereoscopic disparity set to zero. The image presented to the left eye was flickering at 15 Hz while the one presented to the right eye was flickering at 20 Hz.

by the two different flickering frequencies, occurring only when both the frequencies are spatially superimposed. Interestingly, from Fig. 6.2.4, can be noticed that some peaks, as the 20 Hz one, seems to have been suppressed, although its 10 Hz

subharmonic and its 40 Hz harmonic are clearly present.

From the reported graphs, the PSD of the SSVEP responses elicited by the three different stimuli showed to the subject are clearly different and thus a classification seems to be feasible.

The same experiment was performed using also two different flickering frequencies corresponding to 12 Hz and 15 Hz and similar results can be appreciated.

In Fig. 6.2.5 is reported the PSD of the EEG signal acquired by the Oz electrode while the subject was attending the dichoptic flickering white patch, where the image presented to the left eye was flickering at 12 Hz while the one presented to the right eye was flickering at 15 Hz. Also in this case non linear interactions can be appreciated and interestingly a new peak at 9 Hz seems to arise with an higher harmonic at 18 Hz.

This second experiment, for both the frequencies couples, was performed on two different subjects obtaining similar results; the frequencies of the peaks arising from the dichoptic stimulation are the same, although their relative intensity varies between subjects.

6.2.5 CONCLUSION AND FUTURE WORKS

Form the first experiment discussed in this section can be inferred that the stereoscopic disparity of a flickering stimulus do not interfere significantly with the elicited SSVEP response, at least for the used kind of stimulus which was a commonly used unpatterned patch. Consequently, prospective SSVEP based BCI applications in the context of stereoscopic VR environments, could implement integrated flickering stimuli also in objects having a non-zero stereoscopic disparity, without affecting the SSVEP response strength.

Moreover, the used display implements an active stereoscopic technology where the user has to wear shuttering glasses, thus the first experiment, as well as the graphs reported in Fig. 6.2.3, highlight the fact that the synchronization between the screen refresh and the glasses is precise enough to not introduce unwanted flickers at spurious frequencies nor frequency beatings.

A further interesting experiment would be to compare the performance in terms of the elicited SSVEP response between an active stereoscopic display and a pas-

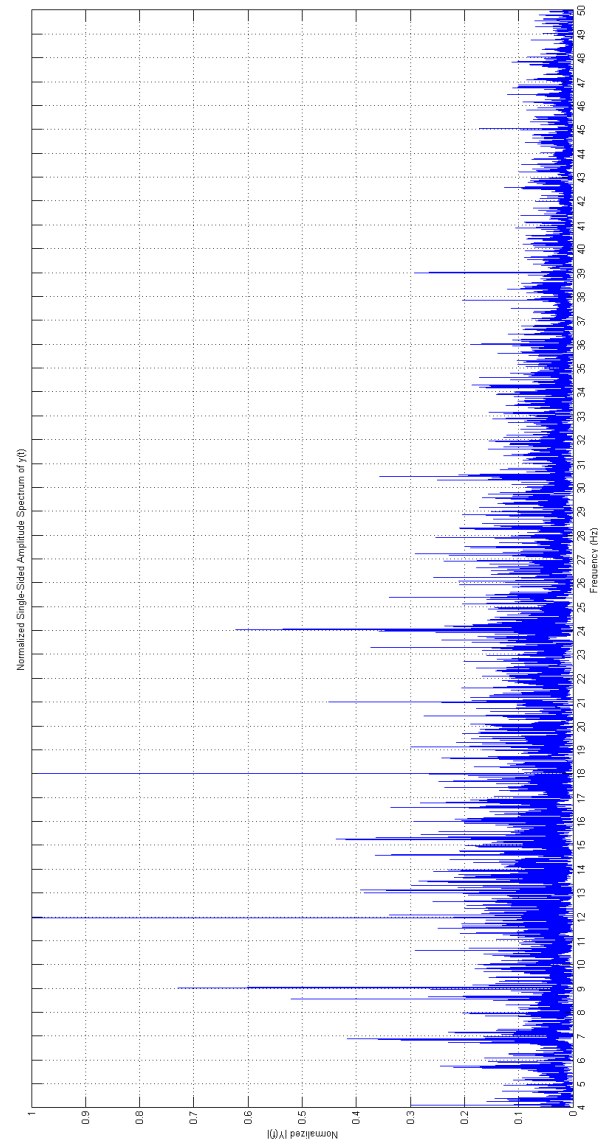


Figure 6.2.5: The normalized PSD of the EEG signal acquired by the Oz electrode while the subject was attending a dichoptic flickering white patch, stereoscopic disparity was set to zero. The image presented to the left eye was flickering at 12 Hz while the one presented to the right eye was flickering at 15 Hz.

sive one. In fact, although higher frequencies seems not to have been introduced in the PSD of the signals, the shuttering glasses are anyway blocking alternatively the light entering in each eye at a frequency of 60 Hz, thus this may affect the duty cycle of the provided stimulus, on its turn affecting the SSVEP response in yet unknown ways.

On the other side, the second experiment, apart from confirming the usability of active stereoscopic displays to provide reliable SSVEP stimuli, it demonstrates also how these devices could be used to provide dichoptic SSVEP stimuli. The reported graphs highlight the possibility of implementing three different targets using only two different flickering frequencies, which is an interesting result considering the limited set of properly displayable frequencies on computer monitor devices.

Moreover, apart from SSVEP based BCIs applications, the dichoptic stimulation used in the second experiment could find applications also in other research fields. In particular, the custom software described in Sec. 5.4.2, being able to independently control the flickering images displayed in the two eye views and their respective flickering frequencies, thanks to commodity stereoscopic display devices, could find applications in all of the research fields where experiments exploiting the SSVEP based binocular rivalry paradigm¹⁰ are commonly performed [1, 188].

6.3 TOWARDS SSVEP BASED *PASSIVE* BCIs

As previously mentioned, BCIs lately gained a lot of attention in the field of VR environments and gaming, both as a mean of new explicit interaction devices, but also, more recently concerning *Passive* BCIs, as a mean of implicit interaction.

The SSVEP response strength and its relatively easy detection lead SSVEP based *Reactive* BCIs to often reach higher performances than BCIs based on other modalities. Moreover the advent of software able to reliably display flickering stimuli on ordinary and stereoscopic displays (as the one implemented in this work) will let be more affordable in the next future to embed SSVEP eliciting stimuli inside VR environments and Computer Games [109].

¹⁰Binocular rivalry is a phenomenon of visual perception in which perception alternates between different images presented to each eye. If the two images are flickering at different frequencies, the amplitude of the SSVEP responses at the two frequencies alternates accordingly to the user's perception.

As introduced in Sec. 4.2.9 the SSVEP response is not only a mechanical reaction to an external stimulus, but it is modulated by the user attention towards the flickering stimulus [87, 183]. Moreover, it seems to be widely accepted that the SSVEP response strength is modulated also by the semantic content of the flickering stimulus [6, 195], although further research would be needed to examine if the semantic content is just a means to attract more intensely the user attention. What is known is that at least the topography propagation of the response is modulated not only by the attention, but also by the affective content or the kind of emotions elicited of/by the flickering stimuli.

Despite of this, as far as I know, the SSVEP response has never been used to implement *Passive* BCIs, although it has been studied in different research fields as psychology and neuroscience in relation to various cognitive, affective and emotional aspects.

What I mean by *Passive* SSVEP based BCI, is a BCI where a flickering target is showed to the user as for *Reactive* SSVEP based BCIs, but where the user is not aware of the possibility to instruct a command gazing at it. The use of such kind of BCI would be to infer, from the user brain SSVEP responses, information related to its gaze direction, but, more interestingly, about the attention the user is paying to a target or to the emotional arousal elicited by the gazed flickering target.

In this section will be described some experiments where implementations of such a kind of SSVEP based *Passive* BCIs are investigated with the aim to provide new approaches towards implicit interaction with VR environments and Computer Games.

6.3.1 TOWARDS VISUAL ATTENTION TRACKING

In ordinary *Reactive* BCIs the user is aware about the possibility to issue a command to the system according to the gazed flickering target, on the other side, with this experiments my aim is to provide a first proof of concept about the feasibility of using the SSVEP modality to implement a *Passive* BCI where the user is unaware about the purpose of the flickering objects.

Besides the realization of a first *Passive* BCI based on the SSVEP modality, the motivation of this study, developed with the help of Prof. Claudio de'Sperati from

Università Vita-Salute San Raffaele of Milan, is to assess the feasibility of tracking the user's visual attention shifts in a synthetic scene, exploiting a *Passive* SSVEP based BCI approach.

A straightforward way to compute user's gaze point position on a screen is to use eye-tracking devices, however, although eye movements are known to be tightly related to visual attention shifts [63], measuring the gaze point is not a direct measure of visual attention. In fact, although it may seem unnatural, an user may indeed gaze at a particular point while focusing her/his attention on another one.

The use of *Passive* SSVEP based BCIs could lead to the possibility to not only track the user gaze point position, as could do a generic eye-tracking device, but also to the possibility to estimate the user attention towards a particular gaze point. In fact the SSVEP response is known to be modulated by the attention towards the flickering target as discussed in Sec. 4.2.9.

Moreover, as has been recently demonstrated, the SSVEP response could be elicited also covertly attending a flickering target [192], showing that the SSVEP response intensity could actually be modulated by attention shifts, at some extent, independently from the gazed point.

The applications of such a system could be interesting from various point of views in the context of VE environments and Computer Games, for example it could be exploited to better distribute available computational resources to efficiently render a virtual scene, concentrating advanced graphical effects only in the regions attracting the user's attention [72].

Otherwise, in general, it could be used to implement any kind of implicit interaction where information about the scene salience points for a specific user interacting with the system could be exploited.

In the first proposed experiment the possibility to actually track the user's gaze point using a *Passive* SSVEP based BCI approach is initially tested. Assigning to the subjects a simple task, their gaze is shifted across a screen where two backgrounds are flickering at different frequencies. Subjects are unaware of the purpose of the flickers, but in their EEG signals is expected to see an increase in the SSVEP response corresponding to the gazed screen side.

On the other side, in a second experiment, with some refinements and exploiting an eye tracking device, the same approach is tested also for attentional shifts

given a fixed gaze point. In this second experiment a new kind of trial is added, where the subjects are asked to keep their gaze fixed at the center of the screen, while shifting only their attention. A weaker, although similar, SSVEP response increase, as for the first experiment is therefore expected, according to the attended side of the screen.

MATERIALS AND METHODS FIRST EXPERIMENT

The experiment has been performed in the Virtual Theater of the University of Milan, described in Sec. 5.2.1, used in monoscopic modality without using INFITEC filters.

During EEG data acquisition, the 15 users participating in the experiment were simply asked to follow with their gaze a non-flickering small (less than 1° of their visual angle) gray circle displayed over a flickering background.

The background was divided exactly in the middle, splitting it in two areas, a left and a right one, flickering (unpatterned white over black) respectively at 20 Hz and 30 Hz, as sketched in Fig. 6.3.1.

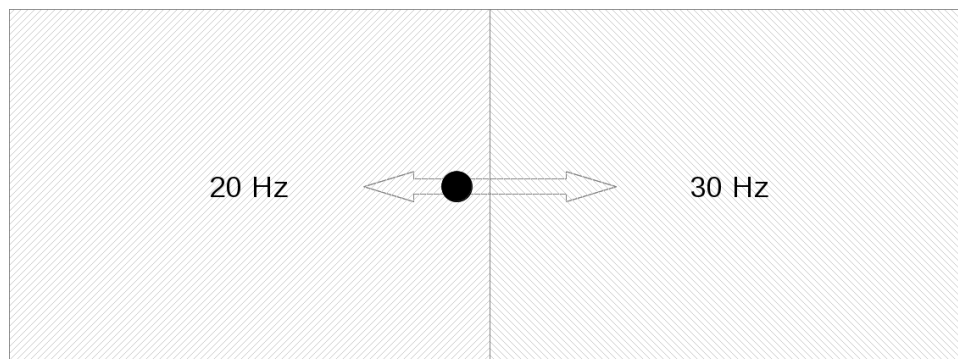


Figure 6.3.1: Sketch of the used stimulus. Arrows were not displayed; they just represent the sinusoidal horizontal oscillation of the circle. Left and right side of the screen where flickering respectively at 20 Hz and 30 Hz.

The circle was moving with a sinusoidally modulated motion along the horizontal direction, starting from the center of the screen and reaching a maximal elongation of 30° to the left and right of the screen center, with respect to the viewer position. This kind of task has been often used to study smooth pursuit eye move-

ments¹¹ induced by following with the gaze a moving object [43]. The purpose of the moving circle has been to provide a task for the users which could assure their gaze point to shift across the two sides of the screen.

To test also the case in the absence of a target, in every trial, for one period of the oscillation, the circle disappeared and the users were instructed to imagine its movement, following it with their gaze also if it was not displayed. Also this task is widely known¹² and its effect concerning the resulting eye movements have often been studied [42, 91].

Every one of the 16 trials consisted in 4 oscillation periods of the circle (one of whose with an “invisible” circle), while the EEG was acquired using the g.Tec MOBIlab+ device described in Sec. 5.1.2, from 4 scalp locations over the visual cortex (POz, CPz, PO7, PO8 positions, according to the extended 10-20 system).

To avoid the influences of stimuli parameters which were not of interest, the flickering frequencies and the moving target direction were swapped at every trial. In half of the trials the stimulus had the 20 Hz flicker on the left (and 30 Hz flicker on the right) and in the other half the opposite. The same for the moving target direction, which in half of the trials was starting its movement to the left and in the other half on the right.

Data was acquired using the OpenVibe framework described in Sec. 5.3 and later saved for off-line analysis. Software triggering was adopted exploiting inter-process communications (as detailed in Sec. 5.3.1) between the custom stimuli presentation software and the *Acquisition Server*, in order to save stimuli events (e.g. the trial starts and the target position) along with the EEG traces.

The SSVEP response was then computed for all the trials for 1 s overlapping windows using the Minimum Energy Combination method and *T* index [53], described in Sec. 4.3.1. Different values of window overlap were tried, but slightly better results were obtained using 0.125 s steps.

¹¹http://en.wikipedia.org/wiki/Smooth_pursuit

¹²http://en.wikipedia.org/wiki/Smooth_pursuit#Smooth_pursuit_in_the_absence_of_a_visual_target

RESULTS FIRST EXPERIMENT

Since the user gaze following the moving circle should oscillate between the two sides of the screen flickering at different frequencies, it is expected to find in the EEG signals of the users an equivalent “oscillating pattern”, showing alternate counter-phase increases in the 20 Hz and 30 Hz SSVEP responses.

As can be seen from Fig. 6.3.2, where data for one subject is reported, experimental results confirm the initial hypothesis.

Actually, the SSVEP response for all the subjects to the 30 Hz flickering frequency revealed to be much weaker than the one to the 20 Hz flicker.

In the T index values corresponding to the 20 Hz stimulation, plotted over time, are indeed clearly visible 4 peaks corresponding to the circle maximal elongation in the 20 Hz flickering region. These peaks are clearly visible for most of the subjects also on a single trial basis, anyhow to have a smoother curve and to also highlight the fact that they were not obtained by chance, in Fig. 6.3.2 is shown a plot obtained by averaging the 16 trials of a single subject. This has been made possible by the high accuracy of the software triggering implemented using the inter-process communication between the stimuli presentation software and the *OpenVibe Acquisition Server*.

To incorporate in a single plot both the contribution of the 20 Hz and 30 Hz SSVEP responses, the plotted values were obtained as a scaled ratio between the two.

The obtained results are highly interesting, since they confirm the feasibility of coarsely tracking the user gaze using the SSVEP response, but, more interestingly they present the possibility of various improvements.

First of all, as already mentioned, they are mainly derived from the 20 Hz response, thus using another flickering frequency instead of the 30 Hz one (a lower one in particular), would lead for sure to better performances.

Moreover, they still do not provide more information than an eye tracker could do, at this stage.

Consequently a second experiment has been planned and performed.

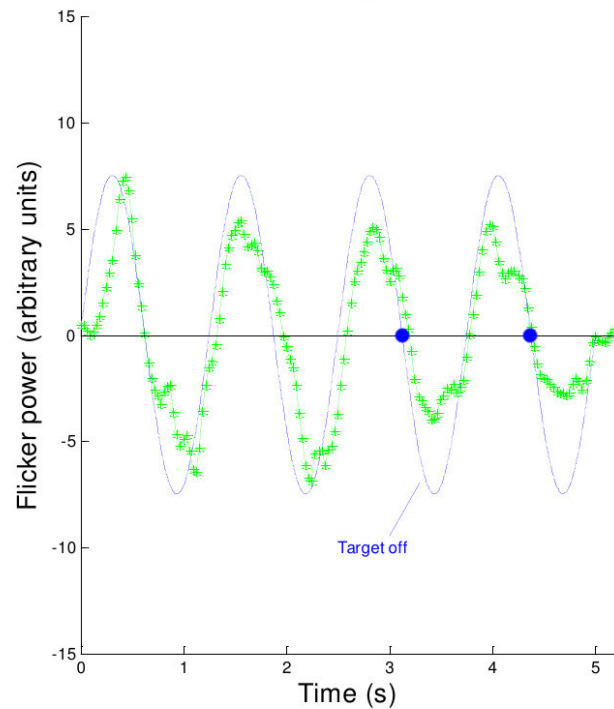


Figure 6.3.2: In this figure is represented the result given by the average across different trials for one subject. The blue solid line represent in arbitrary units the position of the circle on the screen, peaks and valley represent maximal elongations on the left/right of the screen, while every zero-crossing represent the circle crossing the center of the screen. Green dots are measured data points, given in arbitrary units as the average across trials of the ratio between the computed SSVEP responses in the 20 Hz and 30 Hz. The interval between the blue points represent the period in which the circle disappeared and the user was just imagining its presence continuing to follow its imagined movement.

MATERIALS AND METHODS SECOND EXPERIMENT

In this second experiment, some of the issues highlighted by the first results were addressed and moreover different kind of trials were added for further investigations, exploiting also an eye tracking device.

Main aim of this experiment has been to assess the possibility to track attentional shifts independently from the gazed point.

A lower flickering frequency of 15 Hz has been used in place of the 30 Hz one, in order to have a stronger response, thus changing the presented scenario to the one sketched in Fig. 6.3.3.

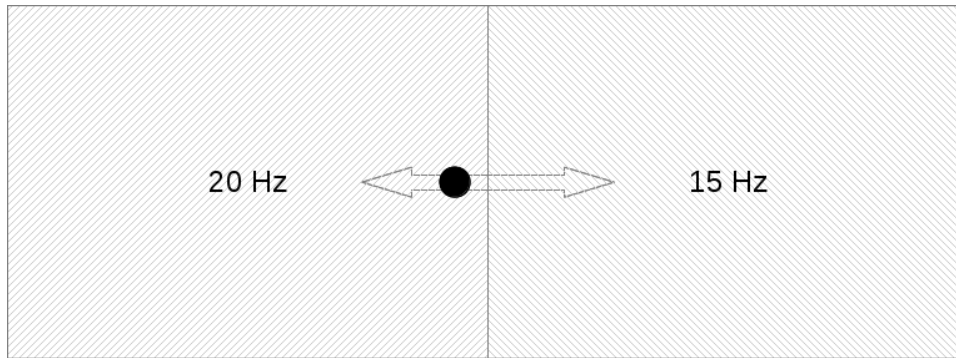


Figure 6.3.3: Sketch of the used stimulus. Arrows were not displayed; they just represent the sinusoidal horizontal oscillation of the circle. Left and right side of the screen where flickering respectively at 20 Hz and 15 Hz.

Moreover, a new kind of task has been introduced in addition to the one described in the first experiment. In this new task, which will be called “fixation task”, the subject is requested to keep her/his gaze fixed at the center of the screen while moving only his/her attention as in the first task he/she was doing with his/her gaze.

To have an objective measure to assess if the subjects correctly performed the task, their eye movements were measured using an infrared oculometer (an eye tracking device).

The eye movements recording assure the possibility to discard the possible trials where the user wrongly shifted the gaze when he/her was supposed to just shift the attention, nevertheless there is no way to assure the user actually correctly shifted the attention.

In Fig. 6.3.4 is reported the hardware configuration utilized in order to record EEG data and eye movements from the subjects performing this experiment.

In order to concurrently record EEG signals and eye movements in a synchronous fashion, the oculometer acquisition software, written in MATLAB has been synchronized to the stimuli presentation software thanks to a TCP socket and thus

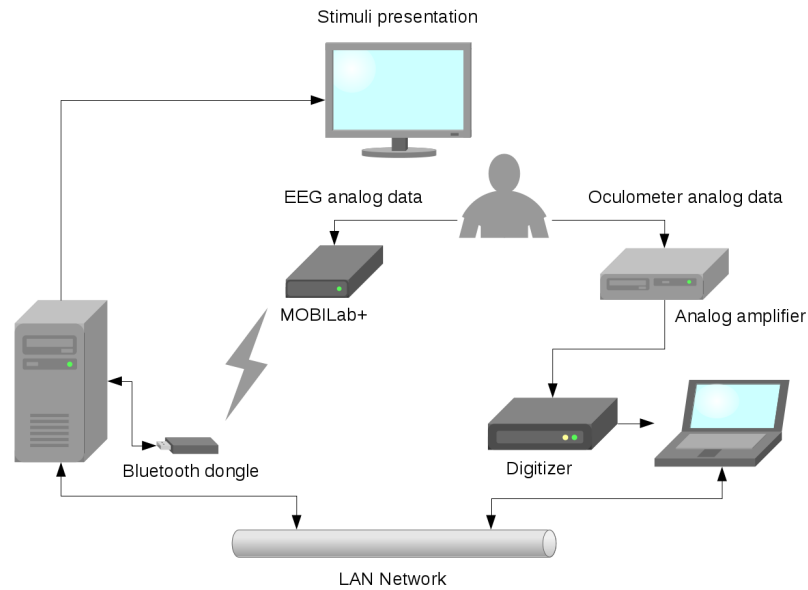


Figure 6.3.4: The used experimental setup, from the hardware point of view, in order to concurrently record EEG and eye tracking data from the user.

synchronized as well to the OpenVibe data *Acquisition Server* as sketched in Fig. 6.3.5.

Thanks to the implemented data acquisition system both the recordings were triggered as soon as the trial started and although two different files were saved in two different machines, as shown in Fig. 6.3.5, they could later be merged while being analyzed off-line.

RESULTS SECOND EXPERIMENT

As can be seen from Fig. 6.3.6 where a single trial result for one subject is reported, the use of 15 Hz flickering frequency in place of the 30 Hz, greatly improved the performance.

Although the result may seems similar to the one reported in Fig. 6.3.2 apart from the oculometer data, it is worth to notice that Fig. 6.3.6 refers to a single trial and thus it is much more interesting.

Moreover, as can be seen from Fig. 6.3.7, where is reported a single trial regard-

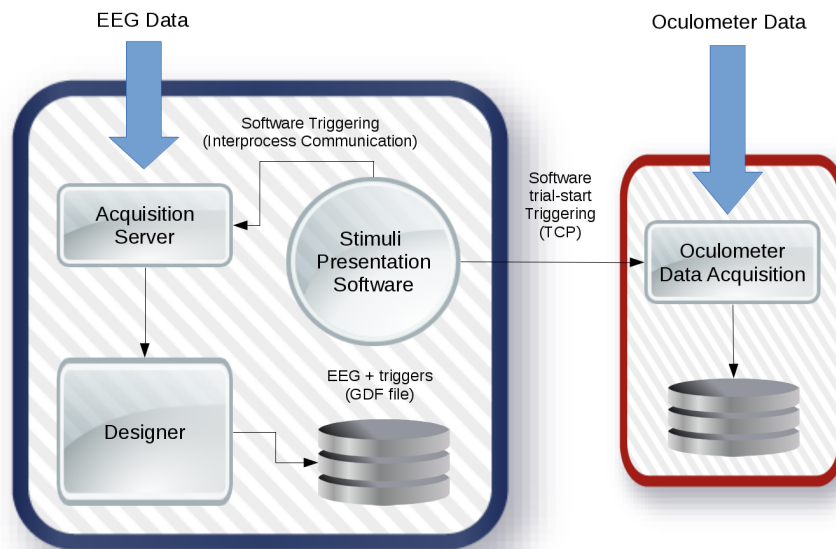


Figure 6.3.5: The used experimental setup, from the software point of view, in order to record EEG and eye tracking data from the user. On the left the computer used to record EEG data, while on the right the laptop used for the oculometer data acquisition.

ing the same subject performing the fixation task, also in this case an oscillation of the SSVEP response ratio following the attentional shift can be appreciated.

CONCLUSION AND FUTURE WORKS

As shown in the previous section, the results of this experiment confirmed the initial hypothesis showing that the use of a *Passive* BCI in order to track the user visual attention is possible.

Interestingly, despite of the fact that the users were concentrating their visual attention on the non-flickering circle, the measured SSVEP response according to the background flickering frequency was strong enough to be clearly detected. Moreover, the same response was detectable also in the absence of the non-flickering circle with similar intensity; the slightly weaker response between the blue points in Fig. 6.3.2 could be explained by the imprecise time alignment between different

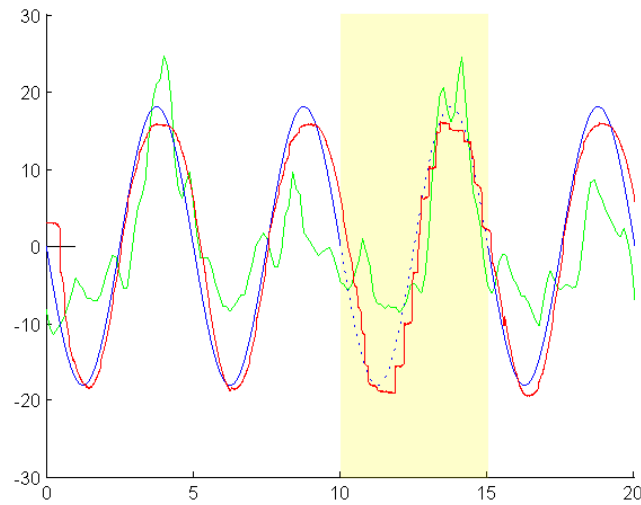


Figure 6.3.6: Single trial result for one subject; seconds on the x-axis, while arbitrary units on the y-axis. The subject was performing the smooth-pursuit task from 0 s to 10 s and 15 s to 20 s, while imagining the moving circle between 10 s to 15 s. The blue solid line represent in arbitrary units the position of the circle on the screen, peaks and valley represent maximal elongations on the left/right of the screen, while every zero-crossing represent the circle crossing the center of the screen. Green line is the measured data, given in arbitrary units as the scaled ratio between the computed SSVEP responses in the 15 Hz and 20 Hz. The red line is the measured eye position along the horizontal direction.

trials, since while imagining the position of the (for that period “invisible”) circle the users could over or under estimate the circle speed differently for each trial. Concerning the single trial plot reported in Fig. 6.3.6 the peaks corresponding to the imagination part are indeed much more pronounced.

The proposed experiment is yet a proof of concept, since only two regions in the screen were used and moreover a particular setup was adopted, using a wide screen able to provide stimuli covering almost the whole visual angle of the user.

The use of more frequencies, e.g. dividing the screen in four quarters, would led to a finer spatial resolution.

Further experiments are for sure worth to be conducted, trying different flickering frequencies, but also different kind of flickering background stimuli.

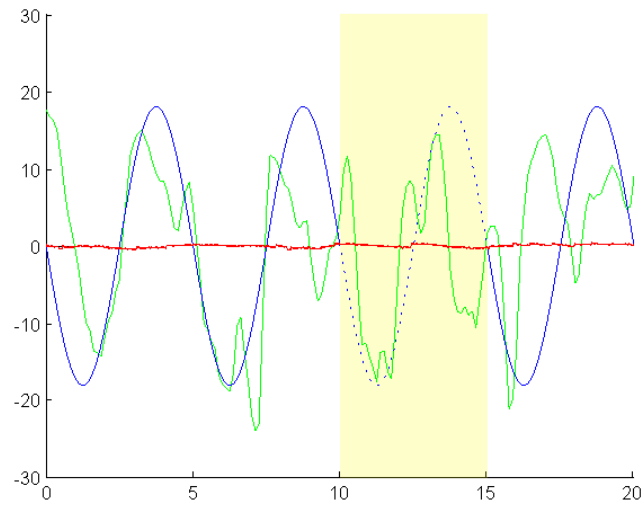


Figure 6.3.7: Single trial result for one subject; seconds on the x-axis, while arbitrary units on the y-axis. The subject was performing the fixation task, while shifting only his attention towards the moving circle from 0 s to 10 s and 15 s to 20 s, while imagining the moving circle between 10 s to 15 s. The blue solid line represent in arbitrary units the position of the circle on the screen, peaks and valley represent maximal elongations on the left/right of the screen, while every zero-crossing represent the circle crossing the center of the screen. Green line is the measured data, given in arbitrary units as the scaled ratio between the computed SSVEP responses in the 15 Hz and 20 Hz. The red line is the measured eye position along the horizontal direction highlighting the fact that the gaze was fixed to the screen center for all the trial length.

6.3.2 TOWARDS FLOW STATE ASSESSMENT

Aim of these experiments is to investigate if from the SSVEP responses of an user it could be possible to infer if he/she is in the flow state while playing a computer game or not.

In recent works, where SSVEP based *Reactive* BCIs where used to implement explicit interaction (e.g. to navigate an avatar in a VR environment), it has been observed an enhancement of the BCI's ITR due to an increased user engagement [109]. An ITR enhancement (given the same user, stimuli and processing algorithms) has to be given by an increase in the SSVEP response strength due to the internal user state. This is coherent also with other works, previously mentioned, strongly

correlating the SSVEP response intensity to the user attention.

Knowing the strict relation between the flow state and the attention modulation/diversion, as mentioned in Sec.3.3.1 and addressed more in depth in [38, 88], my initial hypothesis is that in principle, the entering/exiting in/from the flow channel should change the users' attention level towards the task being carried out. The modulation of the attention should therefore on its turn be able to modulate the SSVEP response elicited by a flickering stimulus, if the flickering stimulus is in some way correlated to the task.

Consequently, if the flow state of an user playing with a computer game, is able to indirectly modulate the SSVEP response strength, it should also be possible to measure the SSVEP response in order to extrapolate an estimation of the user state.

In contrast to the approach followed in [129] and in [9], the proposed one would not be able to discriminate between different game states as *anxiety* and *boredom*, but in principle should be less influenced by other uncorrelated brain activity. Actually it will be influenced by the stimulus characteristics, but at least the possible influences are partially known, as detailed in Chap. 4.

In the proposed experiments the main idea is to present to the user the same flickering stimulus, while changing only the game settings in order to possibly modify the user game related state, to later find a correlation between the SSVEP response modulation and the subjective evaluation of the user state of flow.

MATERIAL AND METHODS FIRST EXPERIMENT

As a first experiment in this direction, I implemented a very simple game where few objects are present in the scene, in order to be able to easily control all the stimuli parameters. As shown in Fig. 6.3.8 the implemented game is a kind of simplified version of the "asteroids game", written in OpenGL language. In my implementation, a white triangle in the center of the screen represents a spacecraft able to shot to a target represented by a white circle. The user can rotate the spacecraft using the two shift-keys on the keyboard to take aim at the target and then shoot using the space-key. Whenever a target is fired, it disappears and another one immediately appears in a quasi-random position on the screen.

To be able to influence the user engagement while playing, the angular speed of

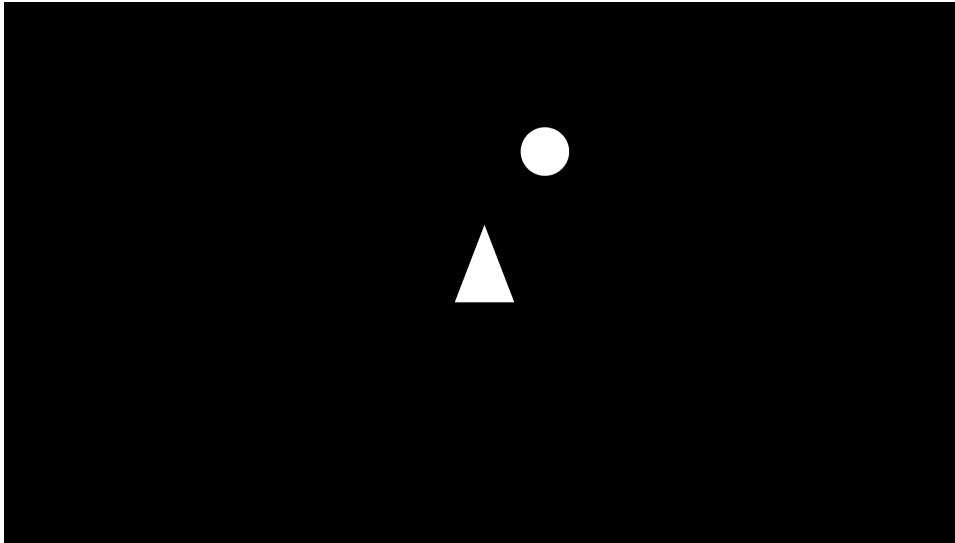


Figure 6.3.8: First game scenario. In this scenario all the visualized objects (spacecraft and target) are flickering in phase at the same frequency.

the spacecraft can be changed in order to let it more or less responsive to the shift-keys. Too low angular speed may let the game be very boring, while too high angular speed may let the game to be frustrating, since the spacecraft became hardly controllable to take aim at the target.

Using the same OpenGL code presented in Sec. 5.4.2 to present reliable SSVEP eliciting stimuli, during the game time, both the spacecraft and the target were flickering at 15 Hz and consequently, being the only objects in the scene, the user (if playing) had to be gazing at one of the two for all the time. This set-up grants a constant flickering stimuli independent from the spacecraft angular speed and from the number of hit/missed targets.

The frequency of 15 Hz has been chosen according to the experiments conducted in [55], since (despite of the SSVEP amplitude distribution showed in Fig. 4.2.1), the SSVEP response seems to be more easily detectable in the 13 Hz to 20 Hz region, probably due to the high natural background brain activity present at lower frequencies. In particular, for this experiment, I choose to avoid the 8 Hz to 13 Hz alpha band, since an high alpha activity is correlated with idleness of the visual cortex, as mentioned in Sec. 2.1.4, thus I preferred to work with a frequency that should not be linked to other visual cortex functions, to avoid a SSVEP response

modulated by even more parameters.

The chosen color for all the objects is white since as found in [24] it should elicit the strongest SSVEP response.

Every experimental trial consisted in 5 minutes of game play, while the EEG was acquired using the g.Tec MOBILab+ device, from 4 scalp locations over the visual cortex (POz, Oz, O₁, O₂ positions according to the extended 10-20 system). A trial could be a “slow” one, designed to induce *boredom* in the user lowering the spacecraft rotation speed, or a “regular” one, designed to be enjoyable and engage the user (as far as possible with a so simple game).

The SSVEP response was computed off-line for all the trials for 1 s non overlapping windows using the Minimum Energy Combination method and *T* index [53], described in Sec. 4.3.1.

RESULTS OF THE FIRST EXPERIMENT

Interestingly, from preliminary results given by experiments conducted on one subject playing 4 trials (2 *slow* ones and 2 *regular* ones), the SSVEP response seems to be correlated to the spacecraft speed, although with opposite sign to what was expected.

In other words, for all the trials, the average SSVEP response over the 5 minutes of game play is higher for *slow* trials and lower for *regular* ones. Moreover, also the standard deviation of the computed values seems to change accordingly to the game settings.

	Regular trials		Slow trials	
	Mean	Std	Mean	Std
Subject 1	1.70	0.74	2.06	1.27
	1.70	0.69	2.14	1.11

Table 6.3.1: Results from the first preliminary experiment performed by one subject. The reported mean is across 300 values, computed using the Minimum Energy Combination algorithm, one for each 1 s window of the 5 minutes of game play.

From the preliminary results the SSVEP response of the subject, on average,

is higher while experiencing *boredom* and lower when engaged in the game. At first this may seem in contrast to the results mentioned in the introduction of this section and also to the results presented in [109].

Despite of this initial guess, it has to be highlighted the fact that in SSVEP based *Reactive* BCIs, the user knows that gazing at a flickering target will produce an action. Moreover in that context, the user can easily learn that the amount of attention paid to the flickering stimulus is correlated to the successful instruction of a command. Consequently for *Reactive* BCIs, the reported increase of the ITR in correlation with an increase of the user engagement, is probably related to the user will to issue the right command. More the user is engaged, more is the user will to succeed in the assigned task and thus more is the attention paid to the flickering targets, leading to an higher SSVEP response and an higher ITR.

In the context of the performed experiment, the user is not only unaware of the function of the flickering stimulus, but is probably distracted by it, from his/her task which is to shoot to the targets.

The obtained results can be therefore interpreted as the fact that the user recognizes the flickering stimulus as uncorrelated to his/her goal and therefore, although forced to look at it, he/her ignores it while engaged in the task. On the other hand, while bored, or anyway, idling, waiting for the spacecraft to move to the desired position, the user is more prone to be distracted by the flickering target.

This is indeed coherent with the observation reported in [38] about the fact that an user experiencing the state of flow while engaged in a task is more hardly prone to be distracted by stimuli uncorrelated to the task than when not experiencing the state of flow.

This new hypothesis is then the basis for another experiment which is later described, that will also try to address another problem evident from the results reported in Tab. 6.3.1: the elicited average SSVEP response, for all the trials, is quite weak and thus more effective stimuli should be adopted.

MATERIAL AND METHODS SECOND EXPERIMENT

In the second experiment a different approach has been followed. In this scenario the same game is proposed, this time with a non-flickering “spacecraft” and non-

flickering targets, while the same “spacecraft” angular speeds, as for the first experiment, have been adopted to provide “slow” and “fast” trials.

On the other hand, another object has been introduced, called “distractor”, role of which is to try to divert the user attention from the game to itself. During the game play the “distractor” appears every 5 s in a quasi-random position (which is never over/under the target or the spacecraft) and flickers at 15 Hz for 5 s.

In order to be more likely to attract the user attention and to elicit a strong SSVEP response if attended, the “distractor” is a white patch with a “smile icon” picture superimposed, as shown in Fig. 6.3.9. This is due to the fact that previous works reported stronger SSVEP response when using flickering stimuli with happy/angry faces [6].

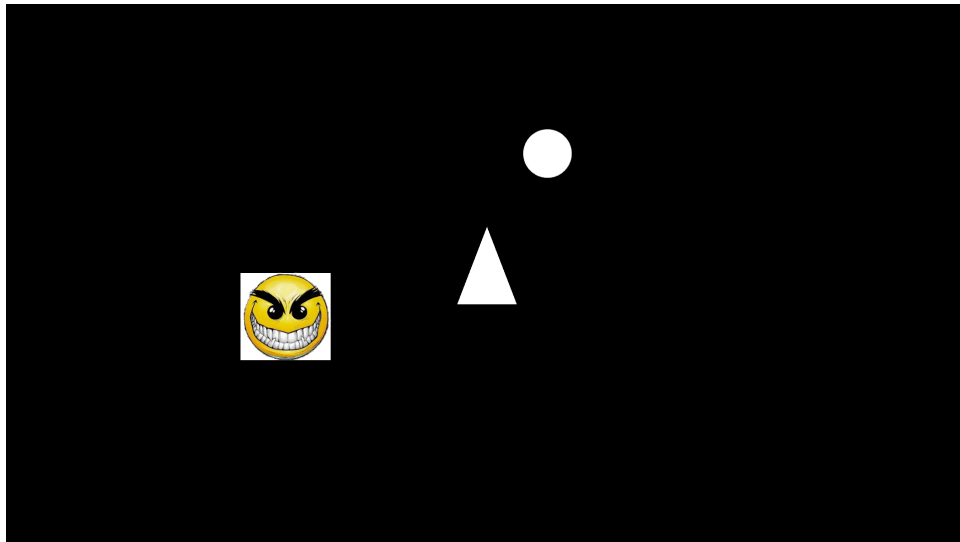


Figure 6.3.9: Second game scenario. In this scenario the spacecraft and the target are not flickering. Every 5 s the distractor (e.g. the smile icon shown on the left) appears in a quasi-random position, flickering at 15 Hz for 5 s and then disappears.

To increase the “temptation” for the user to divert its attention from the game to the “distractor”, the latter do not always present the same picture, but it pseudo-randomly select it from a set of happy and angry smiley which have all approximately the same colors content.

According to the *flow* theory, an user experiencing a state of flow should be less

prone to be distracted, thus in the “slow” trials the average SSVEP response, in the EEG epochs while the “distractor” is present, should be higher than in the “regular” trials.

In other words, according to my hypothesis, the average SSVEP response while the “distractor” is present, should reflect in some sense the “amount” of user’s attention it succeeded to divert from the game play. This value should be high if the user is prone to be distracted and low if the user is not.

EEG signals have been acquired as for the first experiment and later analyzed off-line with the same algorithm, apart from the fact that in this case the SSVEP response has been computed only in the EEG epochs where the “distractor” was present. To this aim the software triggering capabilities of the custom implemented software described in Sec. 5.4.2 have been exploited.

RESULTS SECOND EXPERIMENT

Results from the second experiment, although performed on a single subject, seem to confirm this second hypothesis, since for all the four trials the average SSVEP response while the “distractor” was present, is higher for “slow” trials than for “regular” ones, as shown in Tab. 6.3.2.

	Regular trials		Slow trials	
	Mean	Std	Mean	Std
Subject 1	1.57	0.67	1.87	0.74
	1.30	0.70	2.00	0.86

Table 6.3.2: Results from the second preliminary experiment performed by one subject. The reported mean is across values, computed using the Minimum Energy Combination algorithm, over the EEG epochs where the flickering smiley shown in Fig. 6.3.9 was present.

Unfortunately the SSVEP response is quite weak also in this case. This can be explained by the fact that the user is never really focusing all the attention on the flickering object as he/she would in the case of a regular *Reactive* SSVEP based BCI.

In order to better understand the occurring phenomena, a more in depth study has been performed apart from the considerations about the mean and the standard deviation. The T index has been computed for 1 s overlapping windows for all the signal length and its values have been plotted with respect to the elapsing time, highlighting when the “distractor” was present, for both the “regular” and “slow” trials. Plotted values are respectively reported in Fig. 6.3.10 and Fig. 6.3.11.

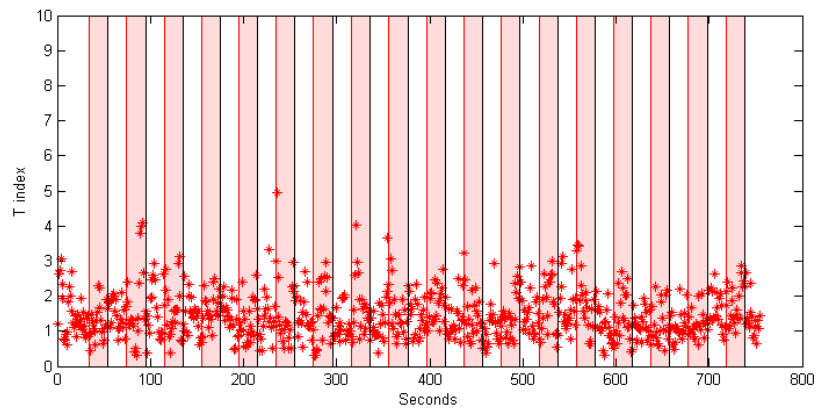


Figure 6.3.10: Plot of the T index value computed for 1 s overlapping windows on the EEG signals acquired from a subject performing one “regular trial”. Red and black vertical lines represent respectively the onset and offset of the “distractor” and thus colored areas represent periods when it was present.

From these graph and in particular from the one relative to the “slow trial”, reported in Fig. 6.3.11, is evident that the mean value is affected by three strong peaks in the T value, corresponding to the first, fourth and tenth apparition of the “distractor”.

Thanks to further experimentation performed with the aid of questionnaires to be filled by the subject after performing the experiment, the mentioned peaks revealed to be caused by an actual shift of the user gaze towards the “distractor”, provoking a stronger SSVEP response. This is indeed a known phenomena involving the different effect on the SSVEP response of covert vs overt attention [192].

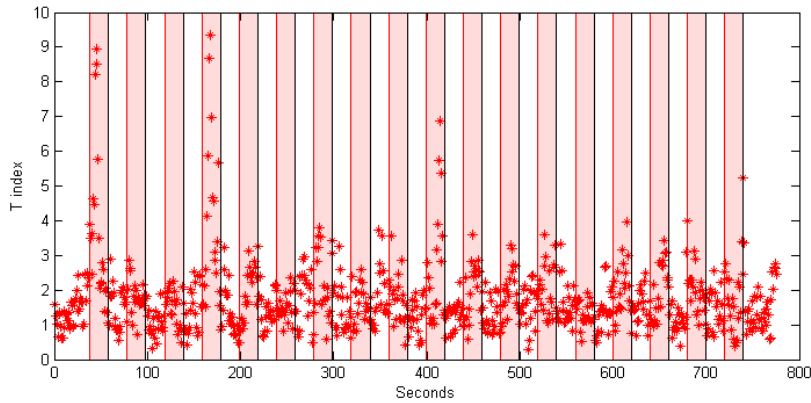


Figure 6.3.11: Plot of the T index value computed for 1 s overlapping windows on the EEG signals acquired from a subject performing one “slow trial”. Red and black vertical lines represent respectively the onset and offset of the “distractor” and thus colored areas represent periods when it was present.

CONCLUSION AND FUTURE WORKS

The results of the second experiment are promising, since they confirm the initial hypothesis and, although performed only on one subject, they at least suggest to continue to experiment using this approach.

Nevertheless, they are not yet satisfactory in terms of practical prospective *Passive* SSVEP based BCIs able to assess the user’s flow state. Different flickering stimuli characteristics could be tried in order to identify an optimal one able to elicit stronger SSVEP responses. In practice, the “distractor” size could be increased, different flickering frequencies could be tried or otherwise, completely different kind of “distractors” could be experimented, for example letting the whole background to flicker, covering a much larger visual angle.

A flickering background in place of the used “distractor” may indeed reduce the difference between the SSVEP response elicited by covertly or overtly attending to it.

In fact, from Fig. 6.3.11 and from the subject’s answers to the questionnaire, is evident that what has been measured is mainly given by the actual eye gaze shift toward the “distractor” and not by the sole attention shift. Consequently the obtained result is not too different from the one obtainable by the use of an eye tracker.

Actually, further experiments with the aid of an eye tracker could be performed in order to assess the influence given by the sole attentional shift with respect to an actual eye shift on the SSVEP response in this context.

Eventually, once identified a final setup, an experiment should be conducted with an higher number of participants which should be asked also to fill a Game Experience Questionnaire (GEQ) [77], in order to assess the elicited game states by the different spacecraft speeds. A correlation between the measured SSVEP response and the experienced game state could then be looked for.

6.3.3 PICTURES EVALUATION

In this Section will be presented an experiment performed in collaboration with Syntyche Gbèhounou, a PhD student from the University of Poitiers (France), where the SSVEP response is investigated for prospective *Passive* BCI applications able to recognize the user's affective reaction to the displayed pictures in a generic VR environment or computer game.

The goal of this experiment is to study a potential relation between SSVEP responses elicited by flickering images and the images features, focusing on the affective content, trying to identify the different contributions given by the objective flickering stimuli characteristics and the one given by the elicited affective state.

In literature there are some works investigating the modulation of the SSVEP response given by pictures containing affective contents, but most of them was performed using the International Affective Picture System (IAPS) [83, 85, 195].

This database contains a particular set of images designed to elicit strong emotions in the observers [96], but with this experiment our aim was to test if the SSVEP response could be used to assess the affective content also regarding natural images which were not specifically created to elicit emotional responses.

This study was indeed performed using a natural and low semantic images database called SENSE (Studies of Emotion on Natural image databaSE) [58]. Where for “low-semantic” is meant that the images do not shock the observers and do not force a strong emotional response.

MATERIAL AND METHODS

The images used during these evaluations were already tagged in the existing database according to the emotions they could elicit and 12 of them were selected according to their valence and arousal values. Images were selected in order to obtain three groups corresponding to positive, neutral and negative valence with different arousal levels.

During the tests, we recorded the EEG of 4 participants while looking at the 12 colored images, one at a time, flickering at a frequency of 10 Hz with 50% duty cycle. The EEG was recorded from 4 electrodes positioned on the occipital area on Pz, POz, PO₃ and PO₄ location according to the extended 10-20 system. The flickering frequency and the electrode locations were selected in conformity with the experiment presented in [83].

All the images were shown to each subject for three times, one time for each of three trials and presented in a pseudo-random order changing for each trial. Every image was displayed for 8 s and then followed by a black screen displayed for 5 s.

The acquired EEG signals were saved to file and later analyzed off-line only were a stimulus was presented, using the Minimum Energy Combination method and the *T* index [53] described in Sec. 4.3.1 and Sec. 4.3.2. Data epochs of 1 s were evaluated and the result averaged over the 8 s of stimulus presence for each of the images.

RESULTS

To study the potential correlation of the SSVEP response strength to various image features, the Pearson's correlation computed by PSPPIRE software¹³ has been used.

At first, the correlation between the computed *T* index values, for each observer, across the different trials, has been evaluated, in order to be sure that there exist a significant modulation of the SSVEP response due to the different pictures content. This first analysis highlighted that a positive correlation exists for all the observers, but only for one observer was significant with a *p* value under 0.05, probably due to the lack of enough data since every image was displayed only three times

¹³<https://www.gnu.org/software/pspp/tour.html>

for each subject.

After a per-subject normalization of the T values to compensate for subjective differences, the correlation has been tested again across the different trials using data from all the observers and the results are reported in Tab. 6.3.3.

Trials		Trial 1	Trial 2	Trial 3
Trial 1	Pearson's r	1	0.62	0.61
	Sig. (bi-var.)		0.00	0.00
	Population	48	48	48
Trial 2	Pearson's r	0.62	1	0.68
	Sig. (bi-var.)	0.00		0.00
	Population	48	48	48
Trial 3	Pearson's r	0.61	0.68	1
	Sig. (bi-var.)	0.00	0.00	
	Population	48	48	48

Table 6.3.3: Results of the Pearson's r correlation test highlighting the correlation between the computed T index value, representing the SSVEP response strength, with the flickering image used.

From the results shown in Tab. 6.3.3 a significant correlation is highlighted with the flickering image, between the values computed in the different trials.

This confirms at first that the computed SSVEP responses are correlated to the pictures showed. After this initial test to assess the meaningfulness of the acquired and then computed data, a correlation between it and the features related to the images has been searched for.

The first hypothesis we tested is the correlation between the SSVEP responses and the arousal of the images. Nevertheless, since other objective features are known to modulate the response, we also looked for a correlation with images features as the average luminance, the average luminance of the different RGB components, as well as an index of their spatial frequency content.

To take into account the spatial frequency content Gabor features energy was computed for each picture using Gabor filters [61] which are directly related to Gabor wavelets. The two-dimensional Gabor filter is defined by the function $g_{\lambda, \theta, \phi}(x, y)$

as the multiplication of a cosine/sine (even/odd) wave with a Gaussian windows, as follows, with $x' = x \cos \Theta + y \sin \Theta$ and $y' = y \cos \Theta - x \sin \Theta$:

$$g_{\lambda, \Theta, \phi}(x, y) = \cos \left(2\pi \frac{x'}{\lambda} + \phi \right) \exp \left(\frac{-(x'^2 + \gamma^2 y'^2)}{2\sigma^2} \right) \quad (6.1)$$

As Gabor features, we considered 12 different angles $\Theta \in [0, \pi]$ every $\frac{\pi}{12}$ and 2 phases $\phi \in \{0, -\frac{\pi}{2}\}$ (0 symmetric case and $-\frac{\pi}{2}$ asymmetric). Consequently 24 different filters were used.

We chose an isotropic Gaussian ($\gamma = 1$) with standard deviation $\sigma = 0,56\lambda$ according to the properties of the visual cortex described in [61]. The energy of Gabor features was then computed as the combination of the results of the 12 filtering for each phase. This value is computed for each pixel and in the case of our test we just considered the average across the pixels and the twelve orientations.

Unfortunately, regarding the correlation between the SSVEP responses and the arousal values, as well as for all the objective features computed, we could not reject the null hypothesis.

CONCLUSION AND FUTURE WORKS

From the obtained results we can confirm that also for the used natural images there is a clear and strong correlation between the pictures and the SSVEP response elicited in the observers.

On the other side, we can not at this stage identify which are the main image features modulating the SSVEP response; a clear statistical significance could be reached performing the experiment on an higher number of subjects and/or using an higher number of images. The correlation with other features should be tried as well, since the most important features could have not been in the tested set.

Due to the kind of our database, it could be interesting also to plan new evaluations with the aid of an “eyetracker” to study during the observation duration the change in the SSVEP response according to the gazed region.

Moreover, using an higher number of EEG electrodes and more complex signal processing techniques, to take into account the SSVEP response propagation from the occipital to the parietal and frontal areas of the cerebral cortex, as proposed in [57], more information about the valence and/or arousal of the emotion

involved in the SSVEP response elicitation could be deduced. It is in fact demonstrated that the modulation of the SSVEP response, due to the user's affect state, changes across different scalp locations in correlation with the arousal and valence of the elicited emotion [85].

Despite of this, from the preliminary results, the contribution of the affective modulation of the SSVEP response, regarding natural images, seems not to be predominant with respect to other objective image features, in contrast to the cases where IAPS images were used.

7

Conclusions

The SSVEP response, have been in this work described in depth with the aim of exploiting it to implement SSVEP based BCIs for VR environments and Computer Game applications. Starting from previous findings, collected and presented in the firsts chapters of this thesis, the needs for the implementation of a SSVEP based BCI have been addressed in conjunction with the reasons motivating the adoption of such systems in the context of VR environments and Computer Games.

This study had to span over several research fields, from the ones setting the basis to understand the mechanism of the SSVEP response, as the basic neurobiology of the human brain and neuro-physiology, through bio-engineering for EEG data acquisition, computer science and signals theory for data synchronization and processing, till positive psychology regarding the game related states.

After this multi-disciplinary review, thanks to the acquired knowledge and given the available hardware, a set of software made up by pre-existing codes and custom developed ones has been proposed as a state-of-the-art for the implementation of a generic SSVEP based BCI to be utilized for the interaction with VR environments

and Computer Games.

Eventually, various experiments have been performed, using the proposed software tools, with different aims, addressing some of the issues which are known to limit the adoption of such BCIs in the context of practical applications for end-users, but also proposing novel applications of the SSVEP modality for implicit interaction.

Using the proposed software bundle, a toy EEG device with a single electrode, available on the market at a price one/two orders of magnitude lower than commonly used professional devices, has been demonstrated to be able to record a SSVEP response accurately enough to provide a simple binary classification using EEG epochs short enough for BCI applications. Moreover, a complete SSVEP based BCI system has been implemented and tested over multiple subjects, demonstrating the validity of the proposed method not only for a simple binary classification, but also for an actual self-paced 3-targets SSVEP based BCI.

Using the same software bundle, a commodity stereoscopic display device has been exploited in order to propose a new solution to overcome the limitation imposed by the small set of flickering frequencies utilizable on ordinary displays to provide reliable flickering stimuli. Moreover, the proposed software, exploiting the same approach could find applications also in other research fields where dichoptic flickering stimulation is needed to implement the binocular rivalry paradigm.

Eventually a novel paradigm for SSVEP based BCIs has been proposed in the context of *Passive* BCIs as a means of implicit interaction. In particular, the use of the SSVEP response in the context of *Passive* BCIs has been demonstrated to be successfully exploited in a proof-of-concept experiment where it has been used to track the subjects' visual attention in a VR environment. Moreover it has been proven its usability, not only to track the user gaze point, but also to assess the user attentional shifts while gazing at a fixed location. This proves the possibility to extract from a *Passive* SSVEP based BCI more information than what could be obtained from an eye tracker device.

In view of the presented literature review, of the proposed experiments and of the produced results, practical applications for SSVEP based BCIs seems to be next to come also for end-users applications, although further research is still needed.

One major issue being still an obstacle for practical applications is given by

the lack of hardware acquisition devices specifically meant for the SSVEP detection able to be easily adopted by end-users. In this respect, lately, various bio-engineering companies started to develop hardware devices aimed to ease of use and in the next future is therefore probable that devices based on dry electrode technologies, specifically meant for SSVEP acquisition, will be presented too.

Apart from commercial companies, also makers communities are growing around the research field of BCIs and interesting low-cost EEG devices are being presented in this period to allow for experimentation also by designers and creatives in general. A very interesting project within this scope is the OpenBCI project ¹.

Moreover, apart from the hardware, also an user friendly software to implement such BCIs is still missing, the OpenVibe project is surely moving in the right direction to fill this gap and in the next years would probably become the most used in the field of VR environments and Computer Games.

Despite of this, in this work, the software to provide the flickering stimuli had to be developed from scratch. In this sense, starting from the proposed solution, a more easily utilizable library should be produced in order to be exploitable in various contexts such as pre-existing game engines. A standard API should be engineered and released in order to be integrated in the engines commonly used to implement VR environments and Computer Games.

These technological advances, from the hardware and software point of views, in conjunction with further studies in the HCI research field should lead in the next few years to practical applications also for end-users. Furthermore, novel advances in the understanding of the brain functions, will probably augment the number of exploitable brain signals in particular also in the context of *Passive* BCIs, where practical applications seems still further to come.

¹<http://www.openbci.com/>

Acknowledgments

First of all, I would like to thank my Family for all the support I received in the last years; I would not been able to reach this goal without it.

I would like to thank Erica for having been everytime by my side, but also for having gone through more than one year without vacations because of this thesis; we will soon go to the seaside, I promise.

I would like to thank my advisor Prof. D. Marini and all the colleagues of the University of Milan, as Cristian, Iuri, Gianfranco, Elif, Saim, Davide, and Alessandro. It has been tough to start this research from scratch, but also with their help I succeeded to reach some of the craved results. Moreover, I have to thank Prof. C. de'Sperati who really helped me a lot and Dr. O. Friman for sharing the code of his Minimum Energy Combination method, which have been used extensively in my work.

I would like to spend a word also to thank all the teachers, professors and mentors who instilled in me a love of science and were fundamental in my path to graduate school, as Prof. Lavarone, Prof. S. Zilio, Prof. G. Zampieri, Dr. D. R. Napoli and all the colleagues from the Legnaro National Laboratories of the INFN.

I would like to thank also all of my friends and in particular the past, present and future students of the Pollaio ² self-managed student lounge of the University of Padua; you are the most beautiful expression of the joy to experiment, learn and share. Soldier on!

Eventually I would like to thank also my thesis reviewers for their valuable comments which I hope to have succeeded to follow. I know my contribution to be just a droplet in this research field, but I really hope this work to be useful at least to avoid to someone else to learn it in a way as hard as I did.

²<http://www.pollaio.org/>

References

- [1] David Alais and Randolph Blake. *Binocular rivalry*. The MIT Press, 2005.
- [2] B.Z. Allison. D5.2: Report about (Re) defining BCIs complete. Technical report, Future BNCI Project, November 2011. URL <http://www.future-bnci.org>.
- [3] F Aloise, F Schettini, P Aricò, F Leotta, S Salinari, D Mattia, F Babiloni, and F Cincotti. P300-based brain-computer interface for environmental control: an asynchronous approach. *Journal of Neural Engineering*, 8(2), 2011.
- [4] Hovagim Bakardjian. *Optimization of Steady-State Visual Responses for robust Brain-Computer Interfaces*. PhD thesis, Tokyo University of Agriculture and Technology, 2011.
- [5] Hovagim Bakardjian, Toshihisa Tanaka, and Andrzej Cichocki. Optimization of SSVEP brain responses with application to eight-command Brain-Computer Interface. *Neuroscience letters*, 469(1):34–38, 2010.
- [6] Hovagim Bakardjian, Toshihisa Tanaka, and Andrzej Cichocki. Emotional faces boost up steady-state visual responses for brain-computer interface. *NeuroReport*, 22(3):121–125, 2011.
- [7] Jessica D Bayliss. Use of the evoked potential P3 component for control in a virtual apartment. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 11(2):113–116, 2003.
- [8] Hans Berger. Über das elektrenkephalogramm des menschen. *European Archives of Psychiatry and Clinical Neuroscience*, 87(1):527–570, 1929.
- [9] R. Berta, F. Bellotti, A. De Gloria, D. Pranantha, and C. Schatten. Electroencephalogram and Physiological Signal Analysis for Assessing Flow in Games. *Computational Intelligence and AI in Games, IEEE Transactions on*, 5(2):164–175, 2013.

- [10] Jordi Bieger and Gary Garcia Molina. Light Stimulation Properties to Influence Brain Activity. Technical report, Philips Research, September 2010.
- [11] Jordi Bieger, Gary Garcia Molina, and Danhua Zhu. Effects of Stimulation Properties in Steady State Visual Evoked Potential Based Brain-Computer Interfaces. In *32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2010.
- [12] Benjamin Blankertz, Michael Tangermann, Carmen Vidaurre, Thorsten Dickhaus, Claudia Sannelli, Florin Popescu, Siamac Fazli, Márton Danóczy, Gabriel Curio, and Klaus-Robert Müller. Detecting Mental States by Machine Learning Techniques: The Berlin Brain-Computer Interface. *Brain-Computer Interfaces*, pages 113–135, 2010.
- [13] Benjamin Blankertz, Steven Lemm, Matthias Treder, Stefan Haufe, and Klaus-Robert Müller. Single-trial analysis and classification of ERP components—a tutorial. *Neuroimage*, 56(2):814–825, 2011.
- [14] Doug A Bowman, Ernst Kruijff, Joseph J LaViola Jr, and Ivan Poupyrev. *3D user interfaces: theory and practice*. Addison-Wesley, 2004.
- [15] Jason Boyd and Yixin Chen. An open source stimulator for SSVEP-based BCIs. In *Proceedings of the 50th Annual Southeast Regional Conference*, pages 124–129. ACM, 2012.
- [16] D. H. Brainard. The Psychophysics Toolbox. *Spatial Vision*, 10:433–436, 1997.
- [17] Clemens Brunner, Giuseppe Andreoni, Lugi Bianchi, Benjamin Blankertz, Christian Breitwieser, Shin’ichiro Kanoh, Christian A Kothe, Anatole Lécuyer, Scott Makeig, Jürgen Mellinger, et al. BCI software platforms. In *Towards Practical Brain-Computer Interfaces*, pages 303–331. Springer, 2013.
- [18] P Brunner, S Joshi, S Briskin, J R Wolpaw, H Bischof, and G Schalk. Does the ‘P300’ speller depend on eye gaze? *Journal of Neural Engineering*, 7(5): 056013, 2010.
- [19] György Buzsáki, Costas A Anastassiou, and Christof Koch. The origin of extracellular fields and currents—EEG, ECoG, LFP and spikes. *Nature Reviews Neuroscience*, 13(6):407–420, 2012.

- [20] Enrico Calore, Raffaella Folgieri, Davide Gadia, and Daniele Marini. Analysis of brain activity and response during monoscopic and stereoscopic visualization. In *Stereoscopic Displays and Applications XXIII*, IS&T/SPIE Electronic Imaging, page 8288oM, 2012.
- [21] Enrico Calore, Cristian Bonanomi, Davide Gadia, and Alessandro Rizzi. Test of an open hardware colorimeter. In *CIE Centenary Conference "Towards a New Century of Light"*, pages 620–627. Commission internationale de l'éclairage, 2013.
- [22] Enrico Calore, Davide Gadia, and Daniele Marini. Eliciting Steady State Visual Evoked Potentials by means of stereoscopic displays. In *Stereoscopic Displays and Applications XXV*, IS&T/SPIE Electronic Imaging, page Submitted, Expected early 2014.
- [23] FW Campbell and L Maffei. Electrophysiological evidence for the existence of orientation and size detectors in the human visual system. *The Journal of Physiology*, 207(3):635, 1970.
- [24] Teng Cao, Feng Wan, Peng Un Mak, Pui-In Mak, Mang I Vai, and Yong Hu. Flashing color on the performance of SSVEP-based brain-computer interfaces. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pages 1819–1822. IEEE, 2012.
- [25] Almudena Capilla, Paula Pazo-Alvarez, Alvaro Darriba, Pablo Campo, and Joachim Gross. Steady-state visual evoked potentials can be explained by temporal superposition of transient event-related responses. *PLoS one*, 6(1):e14543, 2011.
- [26] Hubert Cecotti. Classification of Steady-State Visual Evoked Potentials based on the visual stimuli duty cycle. In *Applied Sciences in Biomedical and Communication Technologies (ISABEL), 2010 3rd International Symposium on*, pages 1–5. IEEE, 2010.
- [27] Hubert Cecotti. A self-paced and calibration-less SSVEP-based brain-computer interface speller. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 18(2):127–133, 2010.
- [28] Hubert Cecotti and Bertrand Rivet. Effect of the visual signal structure on Steady-State Visual Evoked Potentials detection. In *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, pages 657–660. IEEE, 2011.

- [29] Hubert Cecotti, Bertrand Rivet, et al. A solution to solve the dilemma of high frequencies and LCD screen for SSVEP responses. *International Journal of bioelectromagnetism*, 2010.
- [30] Hubert Cecotti, Ivan Volosyak, Axel Graser, et al. Reliable visual stimuli on LCD screens for SSVEP based BCI. In *In Proc. of the 18th European Signal Processing Conference (EUSIPCO-2010)*, 2010.
- [31] G. Chanel, C. Rebetez, M. Bétrancourt, and T. Pun. Emotion Assessment From Physiological Signals for Adaptation of Game Difficulty. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 41(6):1052–1063, 2011.
- [32] Bruce G Charlton. The rise of the boy-genius: Psychological neoteny, science and modern life. *Medical Hypotheses*, 67(4):679–681, 2006.
- [33] M Cheng, X Gao, S Gao, and D Xu. Multiple color stimulus induced steady state visual evoked potentials. In *Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE*, volume 2, pages 1012–1014. IEEE, 2001.
- [34] N. Chumerin, N.V. Manyakov, M. van Vliet, A. Robben, A. Combaz, and M Van Hulle. Steady-State Visual Evoked Potential-Based Computer Gaming on a Consumer-Grade EEG Device. *Computational Intelligence and AI in Games, IEEE Transactions on*, 5(2):100–110, 2013.
- [35] Richard M Copenhaver and Nathan W Perry. Factors affecting visually evoked cortical potentials such as impaired vision of varying etiology. *Investigative Ophthalmology & Visual Science*, 3(6):665–675, 1964.
- [36] Tommaso Costa, Elena Rognoni, and Dario Galati. EEG phase synchronization during emotional response to positive and negative film stimuli. *Neuroscience letters*, 406(3):159–164, 2006.
- [37] D. Coyle, J. Principe, F. Lotte, and A. Nijholt. Guest Editorial: Brain/neuronal - Computer game interfaces and interaction. *Computational Intelligence and AI in Games, IEEE Transactions on*, 5(2):77–81, 2013.
- [38] Mihaly Csikszentmihalyi and Isabella Csikszentmihalyi. *Beyond boredom and anxiety: The experience of play in work and games*. Jossey-Bass San Francisco, 1975.
- [39] Edward Cutrell and Desney Tan. BCI for passive input in HCI. In *Proceedings of CHI*, volume 8, pages 1–3, 2008.

- [40] Fernando Lopes da Silva. EEG: Origin and measurement. In *EEG-fMRI*, pages 19–38. Springer, 2010.
- [41] Jean Decety, Daniela Perani, Marc Jeannerod, Valentino Bettinardi, B Tadar, Roger Woods, John C Mazziotta, and Feruccio Fazio. Mapping motor representations with positron emission tomography. *Nature*, 371: 600–602, 1994.
- [42] Claudio de’Sperati and Heiner Deubel. Mental extrapolation of motion modulates responsiveness to visual stimuli. *Vision Research*, 46(16):2593–2601, 2006.
- [43] Claudio de’Sperati and Elisa Santandrea. Smooth pursuit-like eye movements during mental extrapolation of motion: The facilitatory effect of drowsiness. *Cognitive Brain Research*, 25(1):328–338, 2005.
- [44] Pablo F Diez, Vicente A Mut, Enrique M Avila Perona, and Eric Laciari Leber. Asynchronous BCI control using high-frequency SSVEP. *Journal of neuroengineering and rehabilitation*, 8(1):39, 2011.
- [45] Mayank Dobriyal, Nuri Yilmazer, and Rajab Chaloo. Performance analysis of spectral estimation techniques for steady State Visual Evoked Potentials (SSVEPs) based Brain Computer Interfaces (BCIs). In *Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on*, pages 13–18. IEEE, 2011.
- [46] Günter Edlinger, Clemens Holzner, and Christoph Guger. A Hybrid Brain-Computer Interface for Smart Home Control. In Julie A. Jacko, editor, *Human-Computer Interaction. Interaction Techniques and Environments*, volume 6762 of *Lecture Notes in Computer Science*, pages 417–426. Springer Berlin Heidelberg, 2011.
- [47] Giuseppe Erba. Preventing seizures from “Pocket Monsters” A way to control reflex epilepsy. *Neurology*, 57(10):1747–1748, 2001.
- [48] J Faller, R Leeb, G Pfurtscheller, and R Scherer. Avatar navigation in virtual and augmented reality environments using an SSVEP BCI, ICABB 2010. In *Workshop W1 Brain-Computer Interfacing and Virtual Reality*, 2010.
- [49] Josef Faller, Gernot Müller-Putz, Dieter Schmalstieg, and Gert Pfurtscheller. An application framework for controlling an avatar in a desktop-based virtual environment via a software ssvep brain-computer interface. *Presence: Teleoperators and Virtual Environments*, 19(1):25–34, 2010.

- [50] Owen Falzon, Kenneth Camilleri, and Joseph Muscat. Complex-Valued Spatial Filters for SSVEP-Based BCIs With Phase Coding. *Biomedical Engineering, IEEE Transactions on*, 59(9):2486–2495, 2012.
- [51] Lawrence Ashley Farwell and Emanuel Donchin. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and clinical Neurophysiology*, 70(6):510–523, 1988.
- [52] Robert S. Fisher, Graham Harding, Giuseppe Erba, Gregory L. Barkley, and Arnold Wilkins. Photic- and Pattern-induced Seizures: A Review for the Epilepsy Foundation of America Working Group. *Epilepsia*, 46(9):1426–1441, 2005.
- [53] O. Friman, I. Volosyak, and A. Graser. Multiple Channel Detection of Steady-State Visual Evoked Potentials for Brain-Computer Interfaces. *Biomedical Engineering, IEEE Transactions on*, 54(4):742–750, 2007.
- [54] Davide Gadia, Cristian Bonanomi, Maurizio Rossi, Alessandro Rizzi, and Daniele Marini. Color management and color perception issues in a virtual reality theater. In *Stereoscopic Displays and Applications XIX*, volume 6803S of *In IS&T/SPIE Electronic Imaging*, 2008.
- [55] G. Garcia-Molina and Danhua Zhu. Optimal spatial filtering for the steady state visual evoked potential: BCI application. In *Neural Engineering (NER), 2011 5th International IEEE/EMBS Conference on*, pages 156–160, 2011.
- [56] Gary Garcia-Molina and Danhua Zhu. Phase Detection of Visual Evoked Potentials Applied to Brain Computer Interfacing. In *Towards Practical Brain-Computer Interfaces*, chapter 14, pages 269–280. Springer, 2013.
- [57] Gary Garcia-Molina, Tsvetomira Tsoneva, and Anton Nijholt. Emotional brain–computer interfaces. *International Journal of Autonomous and Adaptive Communications Systems*, 6(1):9–25, 2013.
- [58] S. Gbèhounou, F. Lecellier, and C. Fernandez-Maloigne. Extraction of emotional impact in colour images. In *6th European Conference on Colour in Graphics, Imaging, and Vision 2012, CGIV 2012*, pages 314–319, 2012.
- [59] Marcus Geelnard and Camilla Berglund. *GLFW User Guide*, API version 2.7 edition, September 2010.

- [60] Laurent George, Anatole Lécuyer, et al. An overview of research on 'passive' brain-computer interfaces for implicit human-computer interaction. In *International Conference on Applied Bionics and Biomechanics ICABB 2010-Workshop W1 'Brain-Computer Interfacing and Virtual Reality'*, 2010.
- [61] S.E. Grigorescu, N. Petkov, and P. Kruizinga. Comparison of texture features based on Gabor filters. *Image Processing, IEEE Transactions on*, 11(10):1160–1167, oct 2002. ISSN 1057-7149.
- [62] Christoph Groenegress, Clemens Holzner, Christoph Guger, and Mel Slater. Effects of P300-based BCI use on reported presence in a virtual environment. *Presence: Teleoperators and virtual environments*, 19(1):1–11, 2010.
- [63] Rudolf Groner and Marina T. Groner. Attention and eye movement control: An overview. *European archives of psychiatry and neurological sciences*, 239(1):9–16, 1989.
- [64] g.MOBilab+, *Instruction for use*. g.tec medical engineering GmbH, v3.09a edition, .
- [65] *Basics on Biosignal Measurement with g.MOBilab+*. g.tec medical engineering GmbH, v2.12.00 edition, .
- [66] Christoph Guger, Brendan Z Allison, Bernhard Grosswindhager, Robert Prückl, Christoph Hintermüller, Christoph Kapeller, Markus Bruckner, Gunther Krausz, and Guenter Edlinger. How many people could use an SSVEP BCI? *Frontiers in Neuroscience*, 6(169), 2012.
- [67] Hayrettin Gürkök, Danny Plass-Oude Bos, Michel Obbink, Gido Hakvoort, Christian Mühl, and Anton Nijholt. Towards multiplayer BCI games. In *BioSPlay: Workshop on Multiuser and Social Biosignal Adaptive Games and Playful Applications. Workshop at Fun and Games, Leuven, Belgium*, 2010.
- [68] Greg Hajcak, Annmarie MacNamara, and Doreen M Olvet. Event-related potentials, emotion, and emotion regulation: an integrative review. *Developmental neuropsychology*, 35(2):129–155, 2010.
- [69] Eddie Harmon-Jones, Philip A Gable, and Carly K Peterson. The role of asymmetric frontal cortical activity in emotion-related phenomena: A review and update. *Biological psychology*, 84(3):451–462, 2010.

- [70] Manfred Hartmann and Tilmann Kluge. Phase coherent detection of steady-state evoked potentials: theory and performance analysis. In *Neural Engineering, 2007. CNE'07. 3rd International IEEE/EMBS Conference on*, pages 179–183. IEEE, 2007.
- [71] Dominic Heger, Felix Putze, and Tanja Schultz. Online workload recognition from EEG data during cognitive tests and human-machine interaction. In *Proceedings of the 33rd annual German conference on Advances in artificial intelligence, KI'10*, pages 410–417, 2010. ISBN 3-642-16110-3, 978-3-642-16110-0.
- [72] S. Hillaire, A. Lecuyer, T. Regia-Corte, R. Cozot, J. Royan, and G. Breton. Design and Application of Real-Time Visual Attention Model for the Exploration of 3D Virtual Environments. *Visualization and Computer Graphics, IEEE Transactions on*, 18(3):356–368, 2012. ISSN 1077-2626. doi: 10.1109/TVCG.2011.154.
- [73] Ian P Howard and Brian J Rogers. *Perceiving in Depth, Volume 2: Stereoscopic Vision*. Number 29. Oxford University Press, 2012.
- [74] Gan Huang, Lin Yao, Dingguo Zhang, and Xiangyang Zhu. Effect of duty cycle in different frequency domains on SSVEP based BCI: A preliminary study. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pages 5923–5926, 2012.
- [75] P Husar and G Henning. Bispectrum analysis of visually evoked potentials. *Engineering in Medicine and Biology Magazine, IEEE*, 16(1):57–63, 1997.
- [76] Han-Jeong Hwang, Dong Hwan Kim, Chang-Hee Han, and Chang-Hwan Im. A new dual-frequency stimulation method to increase the number of visual stimuli for multi-class SSVEP-based brain-computer interface (BCI). *Brain Research*, 1515(0):66–77, 2013. ISSN 0006-8993.
- [77] Wijnand IJsselsteijn, Yvonne de Kort, Karolien Poels, Audrius Jurgelionis, and Francesco Bellotti. Characterising and measuring user experiences in digital games. In *International Conference on Advances in Computer Entertainment Technology*, volume 2, page 27, 2007.
- [78] Chuan Jia, Xiaorong Gao, Bo Hong, and Shangkai Gao. Frequency and Phase Mixed Coding in SSVEP-Based Brain-Computer Interface. *Biomedical Engineering, IEEE Transactions on*, 58(1):200–206, 2011.

- [79] Helmut Jorke and Markus Fritz. INFITEC-a new stereoscopic visualisation tool by wavelength multiplex imaging. In *Proceedings of Electronic Displays*, September 2003.
- [80] Eric R Kandel, James H Schwartz, Thomas M Jessell, et al. *Principles of neural science*, volume 4. McGraw-Hill New York, 2000.
- [81] Christoph Kapeller, Christoph Hintermüller, and Christoph Guger. Augmented control of an avatar using an SSVEP based BCI. In *Proceedings of the 3rd Augmented Human International Conference*, page 27. ACM, 2012.
- [82] Kapeller, Christoph and Hintermüller, Christoph and Guger, Christoph. Usability of video-overlaying SSVEP based BCIs. In *Proceedings of the 3rd Augmented Human International Conference*, page 26. ACM, 2012.
- [83] Andreas Keil, Thomas Gruber, MatthiasM. Müller, Stephan Moratti, Margarita Stolarova, MargaretM. Bradley, and PeterJ. Lang. Early modulation of visual perception by emotional arousal: Evidence from steady-state visual evoked brain potentials. *Cognitive, Affective, & Behavioral Neuroscience*, 3(3):195–206, 2003. ISSN 1530-7026. doi: 10.3758/CABN.3.3.195.
- [84] S.P. Kelly, E.C. Lalor, R.B. Reilly, and J.J. Foxe. Visual spatial attention tracking using high-density SSVEP data for independent brain-computer communication. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 13(2):172–178, 2005.
- [85] A.H. Kemp, M.A. Gray, P. Eide, R.B. Silberstein, and P.J. Nathan. Steady-State Visually Evoked Potential Topography during Processing of Emotional Valence in Healthy Subjects. *NeuroImage*, 17(4):1684–1692, 2002. ISSN 1053-8119. doi: 10.1006/nimg.2002.1298.
- [86] Mark J Kilgard. The OpenGL utility toolkit (GLUT) programming interface API version 3, 1996.
- [87] Yee Joon Kim, Marcia Grabowecky, Ken A Paller, Krishnakumar Muthu, and Satoru Suzuki. Attention induces synchronization-based response gain in steady-state visual evoked potentials. *Nature neuroscience*, 10(1):117–125, 2006.
- [88] Martin Klasen, René Weber, Tilo TJ Kircher, Krystyna A Mathiak, and Klaus Mathiak. Neural contributions to flow experience during video game playing. *Social cognitive and affective neuroscience*, 7(4):485–495, 2012.

- [89] Mario Kleiner, David Brainard, Denis Pelli, Allen Ingling, Richard Murray, and Christopher Broussard. What's new in Psychtoolbox-3. *Perception*, 36(14):1–1, 2007.
- [90] Tilmann Kluge and Manfred Hartmann. Phase coherent detection of steady-state evoked potentials: experimental results and application to brain-computer interfaces. In *Neural Engineering, 2007. CNE'07. 3rd International IEEE/EMBS Conference on*, pages 425–429. IEEE, 2007.
- [91] Richard J. Krauzlis. The Control of Voluntary Eye Movements: New Perspectives. *The Neuroscientist*, 11(2):124–137, 2005.
- [92] Dean J Krusienski and Brendan Z Allison. Harmonic coupling of steady-state visual evoked potentials. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 5037–5040, 2008.
- [93] Dean J Krusienski, Eric W Sellers, François Cabestaing, Sabri Bayouth, Dennis J McFarland, Theresa M Vaughan, and Jonathan R Wolpaw. A comparison of classification techniques for the P300 Speller. *Journal of neural engineering*, 3(4):299, 2006.
- [94] Jean-Philippe Lachaux, Eugenio Rodriguez, Jacques Martinerie, and Francisco J. Varela. Measuring phase synchrony in brain signals. *Human Brain Mapping*, 8(4):194–208, 1999. ISSN 1097-0193.
- [95] Edmund C Lalor, Simon P Kelly, Ciarán Finucane, Robert Burke, Ray Smith, Richard B Reilly, and Gary Mcdarby. Steady-state VEP-based brain-computer interface control in an immersive 3D gaming environment. *EURASIP journal on applied signal processing*, 2005:3156–3164, 2005.
- [96] Peter J Lang, Margaret M Bradley, and Bruce N Cuthbert. International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical report, University of Florida, Gainesville, FL, 2008.
- [97] Nilli Lavie, Aleksandra Hirst, Jan W de Fockert, and Essi Viding. Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, 133(3):339, 2004.
- [98] Anatole Lécuyer, Fabien Lotte, Richard B Reilly, Robert Leeb, Michitaka Hirose, and Mel Slater. Brain-computer interfaces, virtual reality, and videogames. *Computer*, 41(10):66–72, 2008.

- [99] Anatole Lecuyer, Laurent George, and Maud Marchal. Toward Adaptive VR Simulators Combining Visual, Haptic, and Brain-Computer Interfaces. *Computer Graphics and Applications, IEEE*, 33(5):18–23, 2013.
- [100] Stephen Lee and John Kruse. Biopotential electrode sensors in ECG/EEG/EMG systems. *Analog Devices*, 2008.
- [101] Robert Leeb. *Brain-Computer Communication: The Motivation, Aim, and Impact of Virtual Feedback*. PhD thesis, Graz University of technology, 2008.
- [102] Robert Leeb, Doron Friedman, Gernot R Müller-Putz, Reinhold Scherer, Mel Slater, and Gert Pfurtscheller. Self-paced (asynchronous) BCI control of a wheelchair in virtual environments: a case study with a tetraplegic. *Computational intelligence and neuroscience*, 2007, 2007.
- [103] Robert Leeb, Reinhold Scherer, Claudia Keinrath, Gert Pfurtscheller, Doron Friedman, Felix Y Lee, Horst Bischof, and Mel Slater. *Combining BCI and Virtual Reality: Scouting Virtual Worlds*, chapter 23, pages 393–407. MIT Press, 2007.
- [104] J. Legeny, R. Viciano-Abad, and A. Lecuyer. Toward Contextual SSVEP-Based BCI Controller: Smart Activation of Stimuli and Control Weighting. *Computational Intelligence and AI in Games, IEEE Transactions on*, 5(2):111–116, 2013.
- [105] Legény Jozef, Abad Raquel Viciano, and Lécuyer Anatole. Navigating in Virtual Worlds Using a Self-Paced SSVEP-Based Brain-Computer Interface with Integrated Stimulation and Real-Time Feedback. *Presence: Teleoperators and Virtual Environments*, 20(6):529–544, 2011. ISSN 1054-7460.
- [106] Yue Liu, Xiao Jiang, Teng Cao, Feng Wan, Peng Un Mak, Pui-In Mak, and Mang I Vai. Implementation of SSVEP based BCI with Emotiv EPOC. In *Virtual Environments Human-Computer Interfaces and Measurement Systems (VECIMS), 2012 IEEE International Conference on*, pages 34–37. IEEE, 2012.
- [107] M.A. Lopez-Gordo, A. Prieto, F. Pelayo, and C. Morillas. Use of Phase in Brain-Computer Interfaces based on Steady-State Visual Evoked Potentials. *Neural Processing Letters*, 32(1):1–9, 2010.

- [108] Fabien Lotte. Brain-computer interfaces for 3D games: hype or hope? In *Proceedings of the 6th International Conference on Foundations of Digital Games*, pages 325–327. ACM, 2011.
- [109] Fabien Lotte, Josef Faller, Christoph Guger, Yann Renard, Gert Pfurtscheller, Anatole Lécuyer, and Robert Leeb. Combining BCI with virtual reality: Towards new applications and improved BCI. In *Towards Practical Brain-Computer Interfaces*, chapter 10, pages 197–220. Springer, 2013.
- [110] An Luo and Thomas J Sullivan. A user-friendly SSVEP-based brain-computer interface using a time-domain classifier. *Journal of neural engineering*, 7(2):026010, 2010.
- [111] Nikolay V Manyakov, Nikolay Chumerin, Arne Robben, Adrien Combaz, Marijn van Vliet, and Marc M Van Hulle. Sampled sinusoidal stimulation profile and multichannel fuzzy logic classification for monitor-based phase-coded SSVEP brain-computer interfacing. *Journal of neural engineering*, 10(3), 2013.
- [112] John H Martin. The collective electrical behavior of cortical neurons: the electroencephalogram and the mechanisms of epilepsy. *Principles of neural science*, pages 777–791, 1991.
- [113] Dennis J. McFarland and Jonathan R. Wolpaw. Brain-computer interfaces for communication and control. *Commun. ACM*, 54(5):60–66, May 2011.
- [114] Dennis J McFarland, William A Sarnacki, Jonathan R Wolpaw, et al. Brain-computer interface (BCI) operation: optimizing information transfer rates. *Biological psychology*, 63(3):237–251, 2003.
- [115] David G Messerschmitt. Autocorrelation matrix eigenvalues and the power spectrum. Technical report, University of California, June 2006.
- [116] V. Mihajlović, G.G. Molina, and J. Peuscher. To what extent can dry and water-based EEG electrodes replace conductive gel ones?: A Steady State Visual Evoked Potential Brain-computer Interface Case Study. In *BIODEVICES 2012 - Proceedings of the International Conference on Biomedical Electronics and Devices*, pages 14–26, 2012.
- [117] Vojkan Mihajlović, Gary Garcia-Molina, and Jan Peuscher. Dry and Water-Based EEG Electrodes in SSVEP-Based BCI Applications. In *Biomedical Engineering Systems and Technologies*, pages 23–40. Springer, 2013.

- [118] G Garcia Molina, D Ibanez, V Mihajlovic, and D Chestakov. Detection of high frequency steady state visual evoked potentials for brain-computer interfaces. In *17th European Signal Processing Conference (EUSIPCO 2009)*, pages 646–650, 2009.
- [119] G.G. Molina, T. Tsoneva, and A. Nijholt. Emotional brain-computer interfaces. In *Affective Computing and Intelligent Interaction and Workshops. ACII 2009. 3rd International Conference on*, pages 1–9, 2009.
- [120] Desmond Morris. *The Naked Ape: A Zoologist's study of the Human Animal*. Cape, 1968.
- [121] C Muhl and Dirk Heylen. Cross-modal elicitation of affective experience. In *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on*, pages 1–12, 2009.
- [122] C. Mühl, H. Gürkök, D. Plass-Oude Bos, M.E. Thurlings, L. Scherffig, M. Duvinage, A.A. Elbakyan, S. Kang, M. Poel, and D.K.J. Heylen. Bacteria Hunt: A multimodal, multiparadigm BCI game. In *Fifth International Summer Workshop on Multimodal Interfaces*, Genua, 2010. University of Genua.
- [123] Christian Mühl. Neurophysiological Assessment of Affective Experience. In *Proceedings of the Doctoral Consortium at the ACII 2009*, pages 89–96, Enschede, 2009.
- [124] Gernot R Müller-Putz, Reinhold Scherer, Christian Brauneis, and Gert Pfurtscheller. Steady-state visual evoked potential (SSVEP)-based communication: impact of harmonic frequency components. *Journal of neural engineering*, 2(4):123, 2005.
- [125] Sungchul Mun, Min-Chul Park, and Sumio Yano. Performance Comparison of a SSVEP BCI Task by Individual Stereoscopic 3D Susceptibility. *International Journal of Human-Computer Interaction*, 29(12):789–797, 2013.
- [126] John Musson and Jiang Li. A comparative survey of PSD estimation methods for EEG signal analysis. In *Student Capstone Conference Proceedings*, April 2010.
- [127] Lennart Nacke. *Affective Ludology: Scientific Measurement of User Experience in Interactive Entertainment*. PhD thesis, Blekinge Institute of Technology School of Computing, 2009.
- [128] Lennart Nacke and Craig A. Lindley. Flow and immersion in first-person shooters: measuring the player's gameplay experience. In *Proceedings of the*

2008 Conference on Future Play: Research, Play, Share, Future Play '08, pages 81–88, 2008.

- [129] Lennart E Nacke, Sophie Stellmach, and Craig A Lindley. Electroencephalographic assessment of player Experience: A Pilot Study in affective ludology. *Simulation & Gaming*, 2010.
- [130] Lennart Erik Nacke, Michael Kalyn, Calvin Lough, and Regan Lee Mandryk. Biofeedback game design: using direct and indirect physiological control to enhance game interaction. In *Proceedings of the 2011 annual conference on Human factors in computing systems*, pages 103–112. ACM, 2011.
- [131] JoachimH Nagel. *Biopotential Amplifiers*, chapter 70. Electrical Engineering Handbook. CRC Press, second edition, dec 2000.
- [132] Jeanne Nakamura and Mihaly Csikszentmihalyi. *The concept of flow*, chapter 7, pages 89–105. Oxford University Press, 2002.
- [133] Kian B Ng, Andrew P Bradley, and Ross Cunningham. Stimulus specificity of a steady-state visual-evoked potential-based brain–computer interface. *Journal of Neural Engineering*, 9(3), 2012.
- [134] Luis Fernando Nicolas-Alonso and Jaime Gomez-Gil. Brain computer interfaces, a review. *Sensors*, 12(2):1211–1279, 2012.
- [135] Miguel Nicolelis. *Beyond Boundaries: The New Neuroscience of Connecting Brains with Machines—and How It Will Change Our Lives*. Times Books, 2011.
- [136] Anton Nijholt, Danny Plass-Oude Bos, and Boris Reuderink. Turning shortcomings into challenges: Brain–computer interfaces for games. *Entertainment Computing*, 1(2):85–94, 2009.
- [137] Guido Nolte, Andreas Ziehe, Vadim V Nikulin, Alois Schlögl, Nicole Krämer, Tom Brismar, and Klaus-Robert Müller. Robustly estimating the flow direction of information in complex physical systems. *Physical Review Letters*, 100(23):234101, 2008.
- [138] J Vernon Odom, Michael Bach, Colin Barber, Mitchell Brigell, Michael F Marmor, Alma Patrizia Tormene, and Graham E Holder. Visual evoked potentials standard (2004). *Documenta ophthalmologica*, 108(2):115–123, 2004.

- [139] Martin Oehler, Peter Neumann, Matthias Becker, Gabriel Curio, and M. Schilling. Extraction of SSVEP signals of a capacitive EEG helmet for Human Machine Interface. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 4495–4498, 2008.
- [140] Piotr Olejniczak. Neurophysiologic basis of EEG. *Journal of clinical neurophysiology*, 23(3):186–189, 2006.
- [141] Rajesh C Panicker, Sadasivan Puthusserypady, and Ying Sun. An asynchronous P300 BCI with SSVEP-based control state detection. *Biomedical Engineering, IEEE Transactions on*, 58(6):1781–1788, 2011.
- [142] Sergio Parini, Luca Maggi, Anna C. Turconi, and Giuseppe Andreoni. A robust and self-paced BCI system based on a four class SSVEP paradigm: algorithms and protocols for a high-transfer-rate direct brain communication. *Intell. Neuroscience*, 2009:2:1–2:11, January 2009. ISSN 1687-5265.
- [143] Maria A Pastor, Julio Artieda, Javier Arbizu, Miguel Valencia, and Jose C Masdeu. Human cerebral activation during steady-state visual-evoked responses. *The journal of neuroscience*, 23(37):11621–11627, 2003.
- [144] Gert Pfurtscheller and Christa Neuper. Motor imagery and direct brain-computer communication. *Proceedings of the IEEE*, 89(7):1123–1134, 2001.
- [145] Rosalind W. Picard, Elias Vyzas, and Jennifer Healey. Toward machine emotional intelligence: Analysis of affective physiological state. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(10):1175–1191, 2001.
- [146] Danny Plass-Oude Bos, Boris Reuderink, Bram Laar, Hayrettin Gürkök, Christian Mühl, Mannes Poel, Anton Nijholt, and Dirk Heylen. Brain-Computer Interfacing and Games. In Desney S. Tan and Anton Nijholt, editors, *Brain-Computer Interfaces, Applying our Minds to Human-Computer Interaction*, Human-Computer Interaction Series, pages 149–178. Springer, 2010.
- [147] Malypoeur Plong, Kai Shen, Marijn van Vliet, Arne Robben, Marc Van Hulle, and Luc Geurts. Accurate Visual Stimulus Presentation Software for EEG experiments. In *Proceedings of the First Asian Conference on Information Systems*, 2012.

- [148] A Plotnikov, N Stakheika, Alessandro De Gloria, C Schatten, Francesco Bellotti, Riccardo Berta, C Fiorini, and F Ansovini. Exploiting real-time EEG analysis for assessing flow in games. In *Advanced Learning Technologies (ICALT), 2012 IEEE 12th International Conference on*, pages 688–689. IEEE, 2012.
- [149] Mannes Poel, Femke Nijboer, Egon L van den Broek, Stephen Fairclough, and Anton Nijholt. Brain computer interfaces as intelligent sensors for enhancing human-computer interaction. In *Proceedings of the 14th ACM international conference on Multimodal interaction*, pages 379–382. ACM, 2012.
- [150] Alan T Pope, Edward H Bogart, and Debbie S Bartolome. Biocybernetic system evaluates indices of operator engagement in automated task. *Biological psychology*, 40(1):187–195, 1995.
- [151] Anne K Porbadnigk, Simon Scholler, Benjamin Blankertz, Arnd Ritz, Matthias Born, Robert Scholl, K Muller, Gabriel Curio, and Matthias S Treder. Revealing the neural response to imperceptible peripheral flicker with machine learning. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 3692–3695. IEEE, 2011.
- [152] Herbert Ramoser, Johannes Muller-Gerking, and Gert Pfurtscheller. Optimal spatial filtering of single trial EEG during imagined hand movement. *Rehabilitation Engineering, IEEE Transactions on*, 8(4):441–446, 2000.
- [153] Pramila Rani, Nilanjan Sarkar, and Changchun Liu. Maintaining optimal challenge in computer games through real-time physiological feedback. In *Proceedings of the 11th International Conference on Human Computer Interaction*, pages 184–192, 2005.
- [154] D Regan. Some characteristics of average steady-state and transient responses evoked by modulated light. *Electroencephalography and clinical neurophysiology*, 20(3):238–248, 1966.
- [155] D Regan. An effect of stimulus colour on average steady-state potentials evoked in man. *Nature*, 210:1056–1057, 1966.
- [156] D. Regan. Evoked potential and psychophysical correlates of changes in stimulus colour and intensity. *Vision Research*, 10(2):163–178, 1970.
- [157] D. Regan. Steady-state evoked potentials. *Journal of the Optical Society of America*, 67(11):1475–1489, 1977.

- [158] D. Regan. Comparison of transient and steady-state methods. *Annals of the New York Academy of Sciences*, 388(1):45–71, 1982.
- [159] D. Regan. *Evoked potentials and color-defined categories*, pages 444–452. Cambridge University Press, 1987.
- [160] D Regan. Some early uses of evoked brain responses in investigations of human visual function. *Vision research*, 49(9):882–897, 2009.
- [161] David Regan. Recent advances in electrical recording from the human brain. *Nature*, 253:401–407, 1975.
- [162] MP Regan and D Regan. Objective investigation of visual function using a nondestructive zoom-FFT technique for evoked potential analysis. *The Canadian journal of neurological sciences. Le journal canadien des sciences neurologiques*, 16(2):168, 1989.
- [163] Yann Renard, Fabien Lotte, Guillaume Gibert, Marco Congedo, Emmanuel Maby, Vincent Delannoy, Olivier Bertrand, and Anatole Lécuyer. Openvibe: An open-source software platform to design, test, and use brain–computer interfaces in real and virtual environments. *Presence: Teleoper. Virtual Environ.*, 19(1):35–53, February 2010. ISSN 1054-7460.
- [164] Fazlollah M Reza. *An introduction to information theory*. Courier Dover Publications, 1961.
- [165] Ricardo Ron-Angevin and Antonio Díaz-Estrella. Brain–computer interface: Changes in performance using virtual reality techniques. *Neuroscience letters*, 449(2):123–127, 2009.
- [166] Jesse Schell. *The Art of Game Design: A book of lenses*. Taylor & Francis US, 2008.
- [167] Reinhold Scherer, Felix Lee, Alois Schlogl, Robert Leeb, Horst Bischof, and Gert Pfurtscheller. Toward self-paced brain–computer communication: navigation through virtual worlds. *Biomedical Engineering, IEEE Transactions on*, 55(2):675–682, 2008.
- [168] Gunar Schirner, Deniz Erdogmus, Kaushik Chowdhury, and Taskin Padir. The Future of Human-in-the-Loop Cyber-Physical Systems. *Computer*, 46(1):36–45, 2013.
- [169] J John Crosley Shaw. *The brain's alpha rhythms and the mind*. Elsevier Amsterdam, 2003.

- [170] John L Sherry. Flow and media enjoyment. *Communication Theory*, 14(4): 328–347, 2004.
- [171] J.L. Shils, M. Litt, B.E. Skolnick, and M.M. Stecker. Bispectral analysis of visual interactions in humans. *Electroencephalography and Clinical Neurophysiology*, 98(2):113–125, 1996.
- [172] Dave Shreiner et al. *OpenGL programming guide: the official guide to learning OpenGL*. Addison-Wesley Professional, 2009.
- [173] I. S S Silva, J.-F. Naviner, and R.C.S. Freire. Compensation of Mismatch Electrodes Impedances in Biopotential Measurement. In *Medical Measurement and Applications, 2006. IEEE International Workshop on*, pages 33–36, 2006.
- [174] Wolf Singer. Consciousness and the binding problem. *Annals of the New York Academy of Sciences*, 929(1):123–146, 2001.
- [175] Wolfgang Skrandies. *Evoked potentials studies of visual information processing*, chapter 4, pages 71–92. Elsevier, 2003.
- [176] Pekka Tallgren, Sampsa Vanhatalo, Kai Kaila, and Juha Voipio. Evaluation of commercially available electrodes and gels for recording of slow EEG potentials. *Clinical Neurophysiology*, 116(4):799–806, 2005.
- [177] Catherine Tallon-Baudry and Olivier Bertrand. Oscillatory gamma activity in humans and its role in object representation. *Trends in cognitive sciences*, 3(4):151–162, 1999.
- [178] Desney S Tan and Anton Nijholt. *Brain-Computer Interfaces: applying our minds to human-computer interaction*. Springer, 2010.
- [179] Russell M Taylor II, Thomas C Hudson, Adam Seeger, Hans Weber, Jeffrey Juliano, and Aron T Helser. 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, 2001.
- [180] Petteri Teikari, Raymond P Najjar, Hemi Malkki, Kenneth Knoblauch, Dominique Dumortier, Claude Gronfier, and Howard M Cooper. An inexpensive Arduino-based LED stimulator system for vision research. *Journal of Neuroscience Methods*, 2012.
- [181] Fei Teng, Yixin Chen, Aik Min Choong, Scott Gustafson, Christopher Reichley, Pamela Lawhead, and Dwight Waddell. Square or sine: finding a

- waveform with high success rate of eliciting SSVEP. *Computational intelligence and neuroscience*, 2011:2, 2011.
- [182] Eoin Thomas, Matthew Dyson, and Maureen Clerc. An analysis of performance evaluation for motor-imagery based BCI. *Journal of neural engineering*, 10(3):031001, 2013.
 - [183] Paolo Toffanin, Ritske de Jong, Addie Johnson, and Sander Martens. Using frequency tagging to quantify attentional deployment in a visual divided attention task. *International Journal of Psychophysiology*, 72(3):289–298, 2009.
 - [184] Hideaki Touyama. *Brain-CAVE Interface Based on Steady-State Visual Evoked Potential*, chapter 27, pages 437–450. InTech, 2008.
 - [185] J. van Erp, F. Lotte, and M. Tangermann. Brain-Computer Interfaces: Beyond Medical Applications. *Computer*, 45(4):26–34, 2012.
 - [186] AC Metting Van Rijn, A Peper, and CA Grimbergen. High-quality recording of bioelectric events. *Medical and Biological Engineering and Computing*, 28(5):389–397, 1990.
 - [187] Francisco Varela, Jean-Philippe Lachaux, Eugenio Rodriguez, and Jacques Martinerie. The brainweb: phase synchronization and large-scale integration. *Nature reviews neuroscience*, 2(4):229–239, 2001.
 - [188] François-Benoît Vialatte, Monique Maurice, Justin Dauwels, and Andrzej Cichocki. Steady-state visually evoked potentials: focus on essential paradigms and future perspectives. *Progress in neurobiology*, 90(4):418–438, 2010.
 - [189] Ivan Volosyak, Hubert Cecotti, and A Graser. Optimal visual stimuli on LCD screens for SSVEP based Brain-Computer Interfaces. In *Neural Engineering, 2009. NER’09. 4th International IEEE/EMBS Conference on*, pages 447–450. IEEE, 2009.
 - [190] Ivan Volosyak, Hubert Cecotti, and Axel Gräser. Impact of frequency selection on LCD screens for SSVEP based brain-computer interfaces. In *Bio-Inspired Systems: Computational and Ambient Intelligence*, pages 706–713. Springer, 2009.
 - [191] Ivan Volosyak, Hubert Cecotti, and Axel Gräser. Steady-state visual evoked potential response-impact of the time segment length. In *Proceedings of the 7th IASTED International Conference*, volume 680, page 284, 2010.

- [192] Sabrina Walter, Clíodhna Quigley, Søren K Andersen, and Matthias M Mueller. Effects of overt and covert attention on the steady-state visual evoked potential. *Neuroscience Letters*, 519(1):37–41, 2012.
- [193] Yijun Wang, Ruiping Wang, Xiaorong Gao, Bo Hong, and Shangkai Gao. A practical VEP-based brain-computer interface. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 14(2):234–240, 2006.
- [194] Yijun Wang, Y-T Wang, and T-P Jung. Visual stimulus design for high-rate SSVEP BCI. *Electronics letters*, 46(15):1057–1058, 2010.
- [195] Wang, Shangfei and Wu, Guobing and Zhu, Yachen. Analysis of Affective Effects on Steady-State Visual Evoked Potential Responses. In Sukhan Lee, Hyungsuck Cho, Kwang-Joon Yoon, and Jangmyung Lee, editors, *Intelligent Autonomous Systems 12*, volume 194 of *Advances in Intelligent Systems and Computing*, pages 757–766. Springer Berlin Heidelberg, 2013.
- [196] René Weber, Ron Tamborini, Amber Westcott-Baker, and Benjamin Kantor. Theorizing Flow and Media Enjoyment as Cognitive Synchronization of Attentional and Reward Networks. *Communication Theory*, 19(4):397–422, 2009.
- [197] Jonathan R Wolpaw, Niels Birbaumer, Dennis J McFarland, Gert Pfurtscheller, Theresa M Vaughan, et al. Brain-computer interfaces for communication and control. *Clinical neurophysiology*, 113(6):767–791, 2002.
- [198] Zhenghua Wu, Yongxiu Lai, Yang Xia, Dan Wu, and Dezhong Yao. Stimulator selection in SSVEP-based BCI. *Medical engineering & physics*, 30(8):1079–1088, 2008.
- [199] Thorsten O Zander and Christian Kothe. Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general. *Journal of Neural Engineering*, 8(2):025005, 2011.
- [200] Thorsten O Zander, Christian Kothe, Sebastian Welke, and Matthias Rötting. Utilizing secondary input from passive brain-computer interfaces for enhancing human-machine interaction. In *Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience*, pages 759–771. Springer, 2009.
- [201] Thorsten O Zander, Christian Kothe, Sabine Jatzev, and Matti Gaertner. *Enhancing human-computer interaction with input from active and passive*

- brain-computer interfaces*, pages 181–199. Human-Computer Interaction Series. Springer, 2010.
- [202] Thorsten Oliver Zander. *Utilizing Brain-Computer Interfaces for Human-Machine Systems*. PhD thesis, Technischen Universität Berlin, 2011.
 - [203] Thorsten Oliver Zander, C Kothe, S Welke, and M Roetting. Enhancing human-machine systems with secondary input from passive brain-computer interfaces. In *Proc of the 4th Int BCI Workshop & Training Course. Graz University of Technology Publishing House, Graz, Austria*, 2008.
 - [204] Thorsten Oliver Zander, Moritz Lehne, Klas Ihme, Sabine Jatzev, Joao Correia, Christian Kothe, Bernd Picht, and Femke Nijboer. A dry EEG-system for scientific research and brain–computer interfaces. *Frontiers in neuroscience*, 5, 2011.
 - [205] Semir Zeki. *A Vision of the Brain*. Oxford Univ Press, 1993.
 - [206] V. Zemon, E. Pinkhasov, and J. Gordon. Electrophysiological tests of neural models: Evidence for nonlinear binocular interactions in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 90 (7):2975–2978, 1993.
 - [207] Danhua Zhu, Jordi Bieger, Gary Garcia Molina, and Ronald M Aarts. A survey of stimulation methods used in SSVEP-based BCIs. *Computational intelligence and neuroscience*, 2010:1, 2010.
 - [208] Danhua Zhu, Gary Garcia Molina, Vojkan Mihajlovic, and Ronald M Aarts. Phase synchrony analysis for SSVEP-based BCIs. In *Computer Engineering and Technology (ICCET), 2010 2nd International Conference on*, volume 2, pages V2–329. IEEE, 2010.
 - [209] Danhua Zhu, Gary Garcia-Molina, Vojkan Mihajlović, and Ronald Aarts. Online BCI implementation of high-frequency phase modulated visual stimuli. *Universal Access in Human-Computer Interaction. Users Diversity*, pages 645–654, 2011.

Colophon

THIS THESIS WAS TYPESET using \LaTeX , originally developed by Leslie Lamport and based on Donald Knuth's \TeX . The body text is set in 11 point Arno Pro, designed by Robert Slimbach in the style of book types from the Aldine Press in Venice, and issued by Adobe in 2007. A template, which can be used to format a PhD thesis with this look and feel, has been released under the permissive MIT (X11) license, and can be found online at github.com/suchow/ or from the author at suchow@post.harvard.edu.