UNIVERSITÀ DEGLI STUDI DI PALERMO
Dottorato di ricerca in ingegneria dell'innovazione tecnologica
Dipartimento di Ingegneria dell'Innovazione Industriale e Digitale
ING-INF/05

INVESTIGATING PERCEPTUAL AND BIOLOGICAL FEEDBACKS IN HUMAN ROBOT INTERACTION

IL DOTTORE
Ing. Salvatore Tramonte

IL COORDINATORE
Ch.mo Prof. Antonio Chella

IL TUTOR
Ch.mo Prof. Antonio Chella

CO TUTOR
Ch.mo Prof. Rosario Sorbello

CICLO XXX.
ANNO CONSEGUIMENTO TITOLO 2018

Il dubbio é l'inizio della conoscenza.
(Cartesio)

Abstract

The Human Robot Interaction (HRI) is a new discipline that has attracted more attention in the last years due to the increasing presence of robots in people's everyday life. It is a field of study dedicated to understanding, designing, and evaluating robotic systems for use by or with humans. Interaction, by definition, requires communication between robots and humans.

A social robot is an autonomous or semi-autonomous robot capable of interacting and communicating with humans or other autonomous physical agents following social behaviours and rules related to its specific role.

In order to interact effectively with the human being, a robot must be able to decode the complex system of human clues during an interaction. This ability is innate in man, as our brains are accustomed to decoding the behavior and attitude of those in front of us through not only actions or words but also through spontaneous signals that show our partners' intentions.

The goal of my research is to to investigate human's state, during the interaction with a robot, by the use of human's biological and perceptual feedbacks.

In particular, I want to provide an answer to the following research question: *is it possible to model human's response, using features that represent human behaviour, during the interaction with a robot?*

To investigate my research question, I considered different HRI paradigms corresponding to the humans' perception of a robot. For each paradigm I derived some features that I explored in a dedicated experimental scenario. The approach used in my investigation has been multidisciplinary as it required to keep into account concepts and techniques borrowed from psychology, biology, cognitive science and medicine.

I considered three state of the art paradigms for human-robot interaction: robot

as *Avatar*, robot as *Teammate* and robot as *Social Mediator*. These paradigms are distinguished by the mental attribution that the person gives to the robot and by the role of the robot within the interaction. In the robot as *Avatar* the robot is seen as a projection of the person in charge of it, to communicate or interact with others. In robot as *Teammate* paradigm, humans and humanoid robot share the workspace and the objects to complete a task. in robot as *Social-Mediator* paradigm, the robot is used for conveying emotion or to support cooperation or learning.

The user's response in the interaction has been modelled deriving different features. In particular, I considered the following features: *Biological feedback* extracted from humans's neurological activity. *Acceptance* assessed by questionaries. *Trust* derived from human's neurological activity. *Honest signals* derived from human's mimicry. *Emotional Response* aquired from an audience during a concert.

To realize the experimental set to test each feature, I used four different robots. The *Telenoid* robot, which is a humanoid minimalistic robot used as avatar and social-mediator. The *Geminoid* robot, which is a human full scale humanoid robot. The *Nao* robot, which is a small size humanoid robot and the *Kuka* KRC 210 which is an industrial robot.

The experimental results of the experiments carried out to validate the research showed the validity of the considered features. In fact, from the results obtained for each study it is possible to derive features to measure humans' feedback. The further evolution of this research will be to equip a robot with a system based on these features to make the robot able to decode human beings mental and perceptual state and to react accordingly to their states and emotions.

Acknowledgments

This thesis is the expression of human and scientific experience born from my encounter with many people and special places.

Along this path I found many advisers who dedicated me their precious time to answer to my questions and to discuss on my beliefs and ideas.

Firstly I wish to thank my mentor prof. Antonio Chella his guidance and knowledge have been a source of inspiration throughout my doctoral experience.

Secondly, I wish to thank prof. Rosario Sorbello, mentor and friend. He was the first to believe in me and in the scientific relevance of my project. His tenacity and guidance helped me even in the most difficult moments.

Special thanks go to Dr. Christopher Guger and all the people who form g.tec medical engineering for their helpfulness, support and commitment.

A special thought goes to my grandmother, I know that you are proud of me.

I would like to thank my mother and my sister for the support and closeness and for having always allowed me to believe and realize all my dreams. Without you, I would never have been who I am and this thesis would never have existed.

Least but not least I wish to thank my girlfriend, Valeria. You were the person who supported me most, facing with me the challenges I faced every day. Thank you for the patience, dedication and love you have given me.

Contents

List of Figures

List of Tables

Glossary

ALS	Amyotrophic Lateral Sclerosys
AT	Attribution
BCI	Brain Computer Interface
BMI	Brain Machine Interface
EEG	Electroencephalography
ERP	Evoked Related Potentials
FFI	Face to Face Interaction
GUI	Graphical User Interface
HC	Healthy Controls
HHI	Human Humanoid Interaction
HHI	Human Humanoid Interaction
HRI	Human Robot Interaction
KE	Knowledge and expectation
LDA	Linear Discriminant Analysis
MI	Motor Imagery
PA	Perceptual Acquaintance
PET	Positron Emission Tomography
RRC	Robot Response Controller
SSVEP	Steady States Visual Evoked Potentials
SVM	Support Vector Machine

Part I

Foreword

Chapter 1

Introduction

1.1 Motivations and Goals

The Human Robot Interaction (HRI) is a new discipline that has attracted more attention in the last years due to the increasing presence of robots in people's everyday life. It is a field of study dedicated to understanding, designing, and evaluating robotic systems for use by or with humans. Interaction, by definition, requires communication between robots and humans.

A social robot is an autonomous or semi-autonomous robot capable of interacting and communicating with humans or other autonomous physical agents following social behaviours and rules related to its specific role.

In order to interact effectively with the human being, a robot must be able to decode the complex system of human clues during an interaction. This ability is innate in man, as our brains are accustomed to decoding the behavior and attitude of those in front of us through not only actions or words but also through spontaneous signals that show our partners' intentions.

According to the state of the art referred in section 2.1 it is possible to ground a Human-Robot Interaction, from a human centered point of view, on different paradigms.

A robot can be perceived as an **avatar**. According to this paradigm the robot becomes an extension of the user himself to extend his physical presence, or accomplish action on his place. This paradigm supposes that the robot is directly teleoperated or supervisioned by the user who gives desires and commands that are translated into robot actions.

Robot can be perceived as **teammate**. Accordingly to this paradigm, robot must be able to interact with humans, showing autonomy and gaining human's *"trust"* to better accomplish jobs or tasks.

Finally a robot can be used as a **social mediator** to support social interaction and to facilitate the cognitive and social development for people with disabilities or to increase the ability of users engaged in the interaction.

The main questions that arises along all these paradigms is: *is it possible to model human's responses during human robot interactions, using new measurable features?* This thesis is focused in finding an answer to this question evaluating the interaction between human and robot intended as an extension of the user, opponent or collaborator and social mediator. The main focus of the research has been held on human feedback using features to create an evaluation system of the quality of interaction with a humanoid robot. To evaluate the interaction, different features have been identified. Those features are a representation of human's feedback during a Human Robot Interaction: **Biological Feedbacks**, **Acceptance**, **Trust** and **Honest Signals** and **Emotional Response**.

The proposed researches question is trasversal to the interaction paradigms referred. I summarize the general plan of the thesis into an high level model, has shown in figure 1.1 where is shown the paradigm analyzed, the features considered and the studies conducted.

The **robot as avatar** paradigm has been evaluated using biological feedback and acceptance. The **Biological Feedback** is derived from the brain features and it's considered as a representation of user's mental state during the interaction with a robot. The biological feedback is defined as function of the user's level of concentration (*Attention*), of his mental activity (*Mental Workload*) and the level of *mental fatigue* during task execution (*Mental Stress*). The test bed for the model has been a robot interaction for people affected by amyotrophic lateral syndrome (ALS), who were requested to interact with a humanoid NAO robot used as avatar of execute actions on their place.

The **Acceptance** has been validated in a human human interaction, where a locked-in ALS patient used a robot controlled through BCI to express himself.

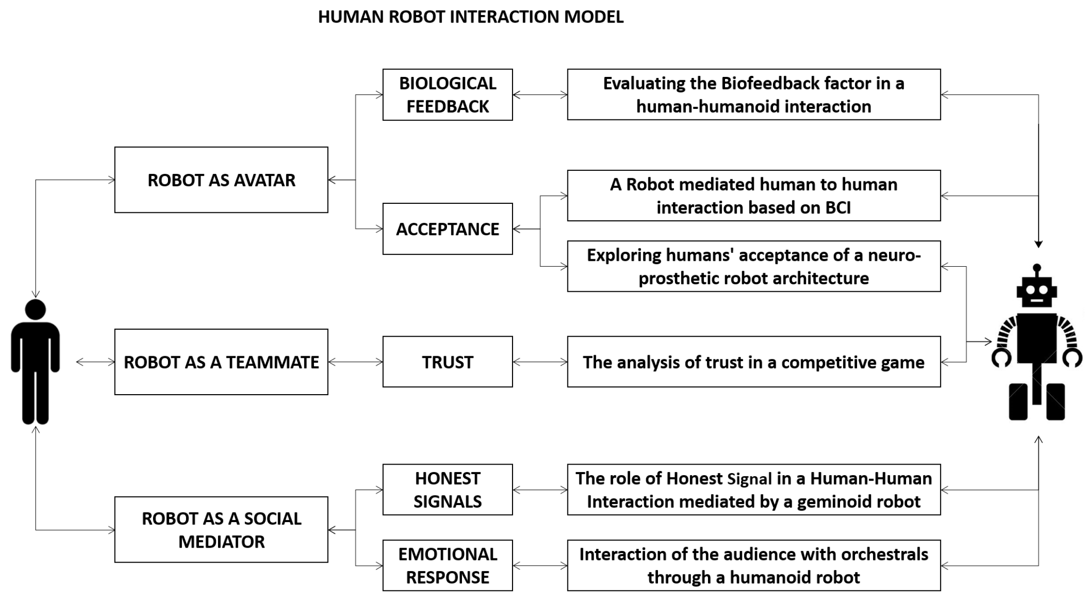


Figure 1.1: The proposed model.

The acceptance of a robot as a **neuroprosthetic extension** has been explored controlling a Kuka robot through Brain Computer Interface, during the Ars Electronica Festival 2017 at Linz, in Austria, for the creation of large artworks. The interaction has been evaluated by visitors to assess users' acceptance.

To explore the **robot as teammate** paradigm, it has been realized a test-bed in which a robot played with a human in a competitive game. In particular a robot "*bargained*" on game rules, to evaluate the connection between biological feedbacks and trust in relation to the robot and to evaluate user's neural response if trustness is broken by violating social rules of the game.

To evaluate the role of the robot as a **social mediator** two studies have been conducted. In the first one, the *Honest Signals* (unconscious signals generated during a human-human interaction mediated by the Geminoid robot) have been evaluated. Experiments have been carried out at the ATR Iroshi Hishiguro Laboratories in Kyoto, Japan, where two subjects were asked to collaborate in solving a problem. In a first phase, the subjects discussed face-to-face. In a second phase, an interaction with the Geminoid, teleoperated from one subject has been conducted.

The second study has been realized in collaboration with the conservatory Vincenzo Bellini of Palermo, Italy. The test-bed has been a live concert in which spectators, through a dedicated mobile application, could modify the performance of the concert through a NAO robot. This study explored the role of the robot as extension of the overall audience.

In conclusion the work of thesis has allowed to explore the human-robot and human-human interaction mediated by the robot by analyzing **biological feedbacks**, **acceptance**, **honest signals**, **trust** and **emotional response**. The goal was to define these features as instruments for evaluating human-robot interaction and human-human robot mediated interaction along the three dimension considered, for instance robot as **avatar**, robot as **teammate** and robot as a **social mediator**.

1.2 Contributions

My research contribution can be summarized as follow:

- I participated in the project "Brain Computer and Robotics in disease management for patients with ALS (Amyotrophic Lateral Sclerosis)". The project was born from the collaboration of RoboticsLab, department of Digital and Industrial Innovation, and the department of Experimental Biomedicine and Neuroscience, University of Palermo the ALS Center of Palermo and ARISLA (Italian Research Foundation for Amyotrophic Lateral Sclerosis). It had as goal investigating the use of humanoid robots, in the management of people suffering from Amyotrophic Lateral Sclerosis, through the use of BCI (Brain Computer Interface) devices. I focused on the concept and the analysis of requirements with patients and shareholders, on the system design and its implementation. Finally I evaluated users acceptance and performances. The studies related to this project are presented in Chapter Chapter 4 and 5.

- Collaboration with g.tec medical engineering, Linz, Austria and the serbian artist Dragan Ilic in the project "A3-k3". The project had as its goal the development of Brain Controlled system to support the artist in the creation of artworks. The project was presented in Linz, Austria at Ars Electronica Festival from 7 to 11 September 2017. In this project I managed the relations between the partners, designed the system and collected the users' data. The study related to this project is presented in Chapter 6.

- Collaboration with the Osaka University and with prof. Iroshi Hishiguro, in the project "Evaluation on Neural feedback in Human-Humanoid Interaction". The goal of the project was the development an architecture to evaluate neural and other physiological measurement to assess user's mental status of engagement, stress, involvement and trust during the interaction with a Geminoid robot. In this project I conceptualized the main experiment, designed the architecture and conducted the experimental procedure. This study is summarized in Chapter 7.

- Collaboration with the Advanced Telecommunications Research Institute International, ATR Kyoto, Japan in the project "Understanding and Transmitting Human Presence". The goal of the project was to evaluate the importance of non-verbal spontaneous gesture during an interaction between a human and a humanoid (Geminoid) robot to asses users' involvement in the interaction. In this project i designed the experiment and evaluated the results. The study related to this project is presented in Chapter 8.

- Collaboration with the musical school "Vincenzo Bellini" of Palermo in the project "Humanoid Robot and Music" for the creation of an innovative performance for robot and musical orchestra. The project was presented with a public performance in September 2016 and June 2017 in occasion of the celebrations for the 400^{th} anniversary of the foundation of the conservatory. I followed the whole project from meeting between partners, till the system design, implementation and data analysis. The study related to this project is presented in Chapter 9.

1.3 Publications

The studies described in this thesis have been published in international journals and international conferences:

1. **Tramonte** S., Sorbello R., Chella A. (2017) Brain Controlled Architecture for Human-Human Interaction Mediated by a Humanoid Robot. International Robotics & Automation Journal 3(5): 00068. DOI: 10.15406/iratj.2017.03.00068. Medcrave.

2. **Tramonte** S., Sorbello R., Giardina M. and Chella A. (2017, July). UnipaBCI a Novel General Software Framework for Brain Computer Interface. *In Conference on Complex, Intelligent, and Software Intensive Systems (pp. 336-348).* Springer, Cham.

3. Giardina M., **Tramonte** S., Gentile V., Vinanzi S., Chella A., Sorce S., & Sorbello R. (2017, July). Conveying Audience Emotions Through Humanoid Robot Gestures to an Orchestra During a Live Musical Exhibition. *In Conference on Complex, Intelligent, and Software Intensive Systems* (pp. 249-261). Springer, Cham.

4. Sorbello R., **Tramonte** S., Giardina M., Cali C., Nishio S., Hishiguro H. and Chella A. (2017). Augmented Embodied Emotions by Geminoid Robot induced by Human Bio-feedback Brain Features in a Musical Experience. *In Conference on Biological Inspired Cognitive Architecture.*. In press.

5. Sorbello R., **Tramonte** S., Giardina M., La Bella V., Spataro R., Allison B. and Chella A. (2017). A Human-Humanoid Interaction through the use of BCI for Locked-In ALS Patients using neuro-biological feedback fusion. *IEEE Transactions on Neural Systems and Rehabilitation Engineering.* In press.

6. Spataro, R., Chella, A., Allison, B., Giardina, M., Sorbello, R., **Tramonte**, S. and La Bella, V. (2017). Reaching and Grasping a Glass of Water by Locked-In ALS Patients through a BCI-Controlled Humanoid Robot. *Frontiers in Human Neuroscience, 11.* DOI:10.3389/fnhum.2017.00068

7. Sorbello R., **Tramonte** S., Cali C., Giardina M., Nishio S., Ishiguro H., and Chella A. (2017, August). An Android Architecture for Bio-inspired Honest Signalling in Human-Humanoid Interaction. In press. *In Conference on Biological Ispired Cognitive Architecture.*

8. Chella A., Sorbello R., Giardina M., **Tramonte** S., Rizzo, G, Orlando L. (2016). Palermo for everyone: towards an inclusive, social town. The project on innovation for inclusive tourism. *Annual Conference on ICT for Smart Cities and Communities.*

9. Spataro, R., Sorbello, R., **Tramonte**, S., Tumminello, G., Giardina, M., Chella, A., and La Bella, V. (2015). Reaching and grasping a glass of water by locked-in ALS patients through a BCI-controlled humanoid robot. *Journal of the Neurological Sciences, 357, e48-e49.*

1.4 Dissertation Outline

The remainder of the dissertation is organized as follows.

- Chapter 2 presents the scientific literature to provide the conceptual grounding and the related works.
- Chapter 3 presents the UnipaBCI framework, a Brain Computer Interface architecture that I designed and implemented to support the latter researches.
- Chapter 4 presents the evaluation the biofeedback factor in a human-humanoid interaction. The biofeedback factor derived from neuro features evaluated in real-time from users during a interaction with the NAO robot.
- Chapter 5 presents the A Robot mediated human to human interaction based on BCI. A Human human brain computer interaction mediated by the Telenoid robot.
- Chapter 6 presents A3-K3 a neuro-prosthetic robot architecture to control a kuka robot using BCI to produce artworks.
- Chapter 7 presents the analysis of trust in a BCI based Human-Humanoid Interaction where a human player plays against a barging NAO robot.
- Chapter 8 presents en evaluation of the role of Honest Signal in a Human-Human Interaction mediated by a Geminoid robot.

- Chapter 9 presents the interaction of the audience with orchestrals through a humanoid robot where a NAO robot is used to express the overall audience's feeling during a live concert.
- Chapter 10 presents the final conclusions and the future direction of this thesis.

Chapter 2

State of the art

In this chapter it is presented an insight on the state of the art relevant for this thesis. First I give an insight on robotics and human-robot interaction providing a description of the paradigms supporting the interaction models proposed in figure 1.1.

Following, I briefly present the Brain Computer Interface to provide an insight on the theory related to the extraction of neurological features.

Finally I give a description of the main features considered for this thesis: *biological feedback*, based on *neurological features*; *honest signals* based on *spontaneous gestures*; *Trust* based on *neurological features* and *emotional response* based on Russel's model of emotion.

2.1 HRI application areas

The Human Robot Interaction (HRI) is a quite new discipline that has attracted more attention in the last years due to the increasing presence of robots in people's everyday life.

It is a field of study dedicated to understanding, designing, and evaluating robotic systems for use by or with humans. Interaction, by definition, requires communication between robots and humans [1], [2]. Accordingly to [3] it is possible to underline four main areas of application:

- Human supervisory control of robots in performance of routine tasks.

- Remote control of robots in hazardous or inaccessible environments.
- Automated vehicles in which a human is a passenger.
- Human-robot social interaction.

The Human supervisory control of robots in performance of routine tasks take in account robots building parts, cutting, handling packages and so on. The HRI explored this dimension, demonstrating techniques to improve human-robot team cooperation during routine tasks [4].

The Remote control of robots in hazardous or inaccessible environments have been explored in many fields along the last years. DARPA (Defense Advanced Research Projects Agency) started in 2012 a robotic competition with the primary technical goal to develop human-supervised ground robots capable of executing complex tasks in dangerous, degraded, human-engineered environments[1]. On the medical field, surgical robots like DaVinci [5] and Zeus [6] are used for operation with the advantage of a minimum invasiveness and the number and types of surgeries being performed with robots is increasing rapidly as more institutions acquire these systems.

The automated vehicles are becoming more popular and mainstream nowadays. Many automobile manufacturers are developing automated veichles to enrich vehicles *skills*, for example break systems, pedestrian avoidance, guidance and navigation integration [7]. Meanwhile Google has produced a self-driven autonomous car guided by artificial intelligence[2].

Besides the other areas, in this thesis the main focus will be held on **Human robot social interaction** which includes a wide broad of robots and sub-areas, including robot to provide entertainment and teaching[8], comfort and assistance for children [9] and elderly [10], autistic [11] and handicapped people [12].

2.2 Social HRI interaction paradigms

People are social beings, and social interaction is relevant in every aspect of our lives, including our interactions with robots. Sociology, the study of society and hu-

[1] https:/www.darpa.mil/program
[2] www.google.com/selfdrivingcar/

man social activity, provides socially-oriented models and methods for how people interact with and understand the world and things [13]. Accordingly to many psychology study, people develop concepts and opinions of subjective "mental states" [14], [15], such as "desiring", "knowing" and "believing" [16].

Some work in the area of HRI considers how robots themselves develop theories of people's minds [17],[18].

Breazeal [19] classifies the modes of interaction with a robotic agent according to four interaction paradigms. Each paradigm is distinguished from the others by the human mental model with which the robot must confront. In the first paradigm, the **robot is seen as a tool**. The human being perceives robots as a tool to achieve a goal. In this paradigm human uses social cues merely as a natural interface for operating (supervising) the robot as a sophisticated tool. This sort of master-slave arrangement does not capture the sense of partnership that we mean when we speak of working "jointly with" humans [20] .

Robot as an extension of a human being. In the second paradigm the robot is a physical extension of a person or part of his body. a clear example is represented by neuro-prosthetic robots [21], [22]. Although the feasibility of such robots is being proved there are still concerns related to ethics [23].

Robot as an avatar. The third paradigm describes the robot as a projection of the person in charge of it, to communicate or interact with others. The robot provides the sense of physical and social presence to the person [24]. The advantage of avatars is that they provide a complete body experience and a physical and social presence [25]. A good example of robot as avatar is represented by the Telenoid R1 which will be described in the following of this chapter.

Hoffman [26] introduces the **Robot as a teammate** paradigm. This paradigm considers a human and an autonomous humanoid robot working together shoulder-to-shoulder, sharing the workspace and the objects required to complete a task. A robotic member of such a team must be able to work towards a shared goal, and be in agreement with the human as to the sequence of actions that will be required to reach that goal, as well as dynamically adjust its plan according to the human's actions . These robots must be equipped with social intelligence so that they can properly interact with humans following social rules and filters [27], [28].

Finally, a robot can be perceived as a **social mediator** for conveying emotions

[29] and to help learning or cooperation [30]. This paradigm is typically used with vulnerable categories like autistic [31], [32] and aged people [25] but they were also used for healthy people who are not affected by any pathology [33].

The Telenoid R1 is a teleoperated android robot designed by the University of Osaka in collaboration with the Advanced Telecommunication Research Institute International [34]. The name was coined by joining the prefix Tele (like telephone and teleprocessing) with the Latin suffix - *oides* (which indicates similarity, as in humanoid).

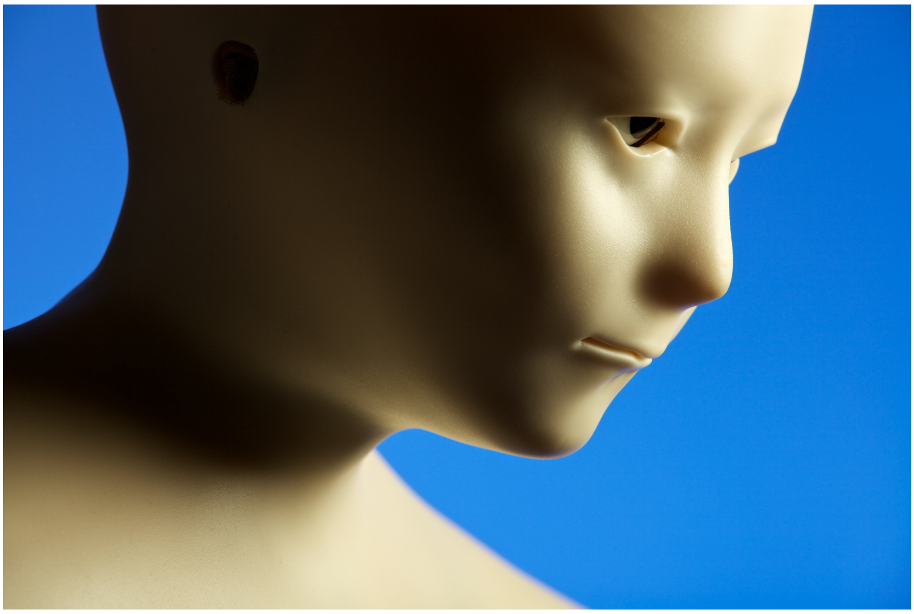


Figure 2.1: The Telenoid R1 Robot.

Telenoid has been designed with the purpose of transferring human presence. Released in August 2011, it is about 80 cm long, weighs about 5 kg and is made of silicone, to provide a pleasant feel to the touch and simulate skin human. It is equipped with 9 servomotors that allow the movement of lips, eyes and eyes, neck and arms.

The teleoperation system is composed by a webcam placed in the center of the head of the Telenoid and two microphones on the sides. A typical Teleoid use case is described by Hannibal in [35].

The choice of providing the Telenoid with a minimalist human design where can be found male and female, young and mature traits provide the ground for a personal interpretation of what Telenoid is: if it reassembles a man or woman, whether young or old, projecting the interlocutor vision on it. Metaphorically speaking, Telenoid becomes a container that man's mind fills, allowing people away from their partners or relatives to feel close to them even if people are far away [35].

2.3 Evaluating Human Robot Interaction

Traditional HCI evaluation is based on measuring the results in accomplishing actions, and goals [36], [37]. This trend also exists in HRI where many studies are devoted to measure performance quality, personal's tactical awareness of the robots' environment and action mistakes [38], [39] and [1].

On the other hand, some researches in HRI specifically targets socially-situated interactions between people and robots, with a particularly strong focus on human's feedback. Much of the research in this area is performed under the title of affective computing, a domain which explores how interaction with an interface influences the emotional state, the feelings, and the satisfaction of the person (Picard, 1999), The approach, based on affective-computing [40] focuses on participants' reactions based on personal experiences (and the context within which it happens). The goal is to describe interaction experience rather than to explicitly measure it starting from the assumption that it is important to accept the complex, unique, and multi-faceted nature of experience which is of difficulty decomposition into constituents [41], [42].

Robots not only provide the cause for explicit social interaction, but are also integrate into everyday environments as social actors. The involvement of social structures in social HRI highlights that, since robots are often viewed as life-like entities (Tisseron 2015), it is possible that person-person structures and norms may manifest between people and robots [27].

2.4 Brain Computer Interfaces

Before the introduction of the features presented in this thesis, it is needed a short degression on Brain Computer Interface to understand some of the concepts and theories treatises in the follow-up.

Brain Computer Interface (BCI), also known as Brain Machine Interface (BMI) is a direct method of communication between a human brain and a computer. This channel of communication is independent from muscular capabilities and it's used by people with sever communicative and motor impairments like people suffering Amyotrophic Lateral Sclerosis (ALS). It measures brain activity associated with the user's intention and translates the recorded brain activity into corresponding control signals for BCI applications [43].

BCI's studies were started by Hans Berger (1873 - 1941) a psychiatry professor who first published in 1929 his study on electroencephalography (EEG), the measurement of electrical potentials generated by synapses with electrodes placed over the scalp [44].

In 1968 Joseph Kamiya demonstrated the possibility to consciously control the alpha rythms. First biofeedback projects were geared to the creation of human-brain interfaces to enanche human cognitive performance for military applications.

In 1973 J. Vidal first coined the term Brain Computer interface, demonstrating how was possible to use evoked potentials and signal processing algorithms to control a cursor in a maze. In the last decades the research on BCI system spread out and many different paradigms and application has been developed.

A first classification of BCI systems considers as discriminating the method of acquisition of the electrical signal: **invasive** or **not invasive** techniques.

In **invasive techniques**, micro-electrodes are implanted directly inside the skull (on the bark or within it); the main advantage consists in obtaining a high signal-to-noise ratio (SNR). This technique brings many questions about long-term effects and is not permitted by the laws of many countries.

In contrast, non-invasive techniques, characterized by electrodes positioned directly on the skin (scalp), have a significantly lower signal quality, in favor of a much higher patient safety and portability that allows a wide use in research and customer applications.

Non-invasive techniques based on EEG have been broadly explored because of their low costs and easiness of montage [45], and for the lack of risks for individuals [46]. Electroencephalogram (EEG) is defined as the measurement of electrical activity of the brain which is revealed as difference of potential between points on the scalp and measured in μV (micro-volt).

In clinical EEG, electrodes are placed following an international montage method called 10-20 system [47] which takes its name from the assumption that the distance between en electrode and the adjacent one is the 10% or the 20% of the total right to left and front to back distance of the skull. Electrodes localization is based on the strict relationship between their position and the underlining cerebral cortex areas. The representation of the 10-20 standard can be found in figure 2.2

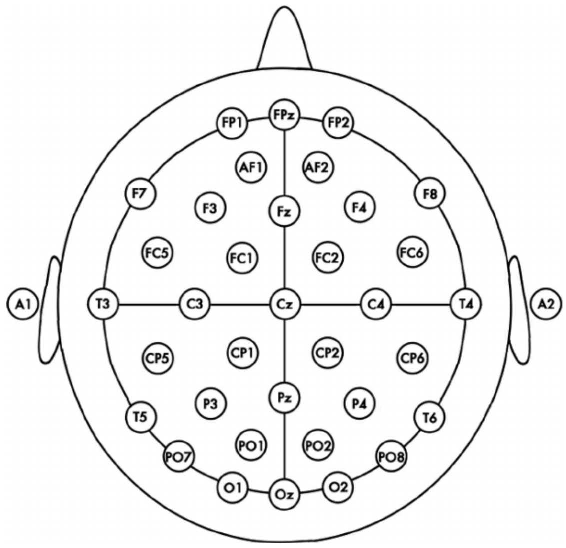


Figure 2.2: The 10-20 standard.

2.4.1 Signal acquisition techniques

Brain, together with the spinal cord, constitutes the central nervous system (CNS) which has the function of interpreting signals, coming both from inside and outside the body processing the responses.

Various techniques based on different physical principles exist to measure brain activity. Between the others let's cite the most used in literature.

Electrocorticography (ECoG) is an invasive technique with excellent spatial and temporal resolution. It uses electrodes placed directly on the exposed surface of the brain to record electrical activity from the cerebral cortex. It has low vulnerability to muscular and environmental artifacts but. Although ECoG systems have been developed over the last ten years also on human individuals, they are not applicable on a large scale to portable BCI systems because of their invasiveness.

Magnetoencephalography has good spatial and temporal resolution, allowing an accurate study of the characteristics of brain signals. This technique, due to the order of magnitude of the magnetic fields to be measured, is not used for the implementation of BCI systems intended for large-scale consumption but is typically used for neuro-physiological investigations.

Positron emission tomography (PET) measures metabolic activity by detecting the activity of radioisotopes placed in the patient. It is not possible to use this technique for a BCI system because of the low spatial resolution and above all for the invasiveness and production cost of the radio-pharmaceutical.

Functional Magnetic Resonance Imaging (FMRI) is a non-invasive imaging technique that allows researchers to detect information about brain metabolism using the BOLD signal (*Blood Oxygen Level Dependent*), dependent on the degree of variable oxygenation of different areas of the brain.

Near infrared spectroscopy (NIRS: Near-Infra Red Spectroscopy) is a non-invasive real-time diagnostic technique, capable of measuring tissue oxygenation using relatively low cost portable instrumentation. This technique is likely to be used in the future for the development of BCI applications.

Electroencephalography (EEG) is a technique based on the assumption that the neurons communicate with each other through electrical impulses. Placing electrodes on the scalp,it is possible to measure the amplitude of electric impulse. Frequency range of a normal brain signal is 1 Hz-100 Hz. Generally, the signal suffers from poor spatial resolution and low signal-to-noise ratio (SNR) since the evoked response to a stimulus is drawn in the background activity.

While recording a signal, various artifacts are combined with the information signals [48]. Different kinds of artifacts that affect the signal are blinking of eyes during signal acquisition procedure, muscular activities, and activities happening in the background. Despite poor spatial resolution, EEGs have excellent temporal resolution of less than a millisecond.

2.4.2 Comparison between Brain Computer Interface techniques

In table 2.1 it is provided a comparison between several of the major recording techniques along 4 principal dimension: *Invasiveness*, *Spatial resolution*, *Temporal Resolution* ans *Cost*. EEG technique is compared with the other classes of techniques that are considered: invasive microelectrode measures (single-unit, multiunit, and local field potential recordings) and hemodynamic measures (PET and fMRI). ERPs are grouped with event-related magnetic fields (ERMFs), which are the magnetic counterpart of ERPs and are extracted from the MEG

Invasiveness. Microelectrodes require brain implant directly in the brain. This led to a small number of individuals who are having electrodes implanted for medical reasons. PET experiments expose users to radiation and to limit an excessive exposure, only a small amount of condition can be tested. In contrast there are no restriction on the amount of ERP or fMRI data that can be acquired from each subject.

Spatial and temporal resolution. Electromagnetic measures have a temporal resolution of 1 ms under optimal condition and an undefined spatial resolution because there are because there are infinitely many internal ERP generator configurations that can explain a given pattern of ERP data. Hemodynamic measures have a limited temporal resolution of hundreds milliseconds and a spatial resolution in the millimeter range. Microelectrode measures have the best temporal and spatial resolutions.

Cost. The ERP technique is quite inexpensive compared to other techniques. An average EEG device costs thousands euros like g.usbAmp, but many commercial solutions are coming on the market at very competitive prices such as the emotiv Epoch+ amplifier that is sold for less the one thousand euros. On the other side fMRI is fairly expensive (typically $500euro/$ hour), and PET is expensive, primarily due to the need for radioactive isotopes with short half-lives and medical personnel. Intra-cranial recordings in humans are not extraordinarily expensive but it is very difficult to get access to the patients.

Parameter	Microelectrode Measures	Hemodynamic Measures	Electromagnetic Measures
Invasiveness	Poor	Good (PET) Excellent (fMRI) Good	Excellent
Spatial resolution	Excellent	God	Undefined/ Poor (ERPs) Undefined /Better (ERMFs)
Temporal Resolution	Excellent	Poor	Excellent
Cost	Fairly expensive	Expensive	Inexpensive (ERPs)Expensive (ERMFs)

Table 2.1: BCI recording technique comparison based on *Invasiveness*, *Spatial resolution*, *Temporal Resolution* ans *Cost*.

2.4.3 Electroencephalograpy and Oscillations

Brain is permanently active, both when one is awake and when sleeping and present a constant PSPs pattern variation to internal an external events. Part of this activity is not driven by discrete events and have an oscillatory nature, reflecting feedback loops in the brain and are named *oscillation* or *rhythms*.

The most prominent oscillation is the alpha wave with a frequency of approximately 10 Hz (from 8 to 12 Hz) as described in figure 2.3.

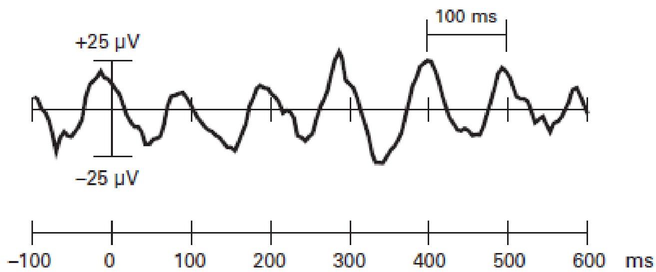


Figure 2.3: Single-trial EEG with large alpha activity. Peaks of the alpha rhythm is separated by approximately 100ms (10 Hz).

Alpha oscillation is most prominent at posterior sites and are particular evident when individual is tired or has his eyes closed. Usually alpha wave is reduced when individual is well-rested or engaged with interesting tasks [49].

Theta Oscillation has been implicated in several brain functions, including sensory processing, memory and control of voluntary movement [50], [51]. It presents a frequency from 4 to 7 Hz. In Gallinat et al. [52] theta oscillation is defined as a synchronous oscillations at distinct frequency ranges are viewed as an important mechanism linking single-neuron activity to behavior and mental disorders. The role of hippocampal theta oscillations (4 - 12 Hz) in mnemonic processes is increasingly targeted in the accumulating body of literature. Theta activity located in the hippocampus area[53], is probably a major operational mode of grouping for neuronal assemblie who cooperates on computational tasks[54].

Beta rhythm is a brain wave with a frequency between 15 and 20 Hz. Beta waves are usually related to normal waking consciousness. Over the motor cortex beta waves are associated with isotonic movements and are suppressed prior to and during movement changes [55]. Bursts of beta activity are associated with a strengthening of sensory feedback in static motor control and reduced when there is movement change [56].

Gamma rhythm has a relatively high frequency (30 - 80 Hz). Gamma oscillation is modulated by sensory input and internal processes such as working memory and attention[57].

In sensory cortex, gamma power increases with sensory drive [58], [59] and with a broad range of cognitive phenomena, including perceptual grouping [60] and attention [61]. At a given recording site, gamma is stronger for some stimuli than others, generally displaying selectivity and a preference similar to that of nearby neuronal spiking activity [62], [63].

All these oscillators, *delta, theta, alpha, beta, gamma*, are usually active in a random way. However, by the application of sensory stimulation, these rythms are coupled and act together in a coherent way. This synchronization and enhancement of EEG activity gives rise to "evoked" or "induced" rhythms. It is possible to derive amplitude and frequency of the individual sine wave by Fourier analysis as shown in figure 2.4.

Evoked potentials (EPs), representing ensembles of neural population responses, were considered to be a result of the transition from a disordered to an ordered state. The compound ERP manifests a superposition of evoked oscillations in the EEG frequencies, ranging from delta to gamma natural frequencies of the brain [64].

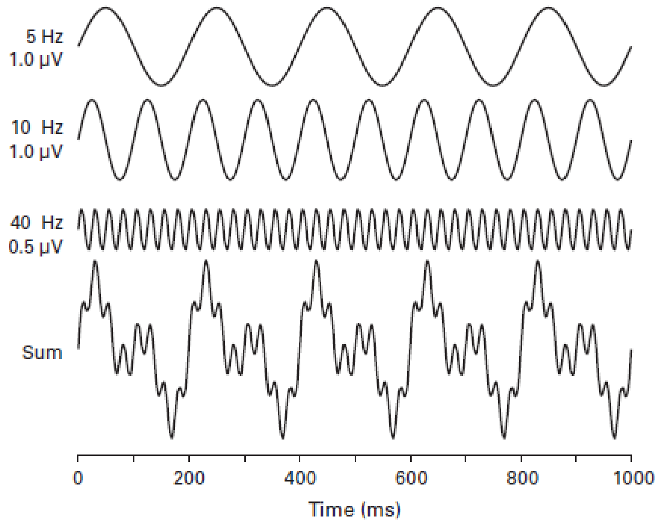


Figure 2.4: An example of mixture of oscillations at different frequencies.Fourier analysis permits to determine amplitude, frequency and phase of each sine wave composing the mixture.

2.4.4 Overview of ERP Component

In the present tesis work signals were acquired using Electroencephalography. It therefore provided an insight on this technique to better address the works described in the subsequent chapters.

From the EEG signal acquired during a task, it is possible to extract **Event Related Potentials** (ERPs) originate as post-synaptic potentials (PSPs), which occur when neurotransmitters bind to receptors, changing the flow of ions across the cell membrane. When PSPs occur in large number of neurons, the resulting potential is conducted at almost speed of light through the brain thus ERPs provide a direct, instantaneous, millisecond-resolution measure of neurotransmission mediated neural activity [65].

ERP components can be either positive or negative since a PSP within a single neuron creates a tiny electrical dipole, namely an oriented flow of current. The polarity of the dipole depends on many factors and it is impossible to draw strong conclusions from ERP polarity. In general, the voltage recorded is positive on one side of the dipole, negative on the other.

2.4.5 The Event Related potentials

ERPs provide an online measure of the processing of stimuli even when there is no behavioral response and are typically used to answer cognitive neuroscience questions. It is possible to detect different peaks, between the others P300 and N400. Typically an ERP component is defined by a letter and a number; the letter is traditionally used to indicate positive or negative peaks, respectively, and the number simply indicates a peak's position, as time latency (expressed in milliseconds), within the waveform. Averaged ERP waveforms consist of a sequence of positive and negative voltage deflections, which are called peaks, waves, or components defined as ERP waveforms, which is the term ERP researchers use to refer to waveforms created by averaging together the averaged waveforms of the person. The use of grand averages [66] masks the variability across people, the variability makes it difficult to see the similarities but may not accurately reflect the pattern of individual results.

In my thesis I derived neurological features from the analysis of the P300 and the N400. **The P300** is composed by different distinguishable ERP components. Squires, and Hillyard (1975) made the first major distinction, identifying a frontally maximal P3a component and a parietally maximal P3b component. Other studies have shown that an unexpected, unusual, or surprising task-irrelevant stimulus within an attended stimulus train will elicit a frontal P3-like response (e.g., Courchesne, Hillyard and Galambos, 1975; Polich and Comerchero, 2003; Soltani and Knight, 2000).

Donchin (1981) proposed that the P3 wave is related to a process he called "context updating" (updating one's representation of the current environment). The mark of the P3 wave is its sensitivity to target probability: As Duncan-Johnson and Donchin (1977) described, P3 amplitude gets larger as target probability gets smaller. Local probability also matters, because the P3 wave elicited by a target becomes larger when it has been preceded by more and more non targets. In addition, P3 amplitude is larger when individuals devote more effort to a task,leading to the proposal that P3 amplitude can be used as a measure of resource allocation as stated by Isreal et al., 1980.

The **N400**, is language-related component as first reported by Kutas and Hillyard (1980). The N400 is negative-going wave that is usually largest over central and parietal electrode sites, with a slightly larger amplitude over the right hemisphere than over the left hemisphere. The N400 is typically seen in response to violations of semantic expectancy and appears to be generated primarily in the left temporal lobe.

The amplitude of the N400 response was most susceptible to manipulation (becoming smaller when factors rendered information more expected and thus easier to process) and most likely to vary with many of the same factors that influence RT measures (see overview in Kutas and Federmeier 2009).

2.4.6 The P300 detection paradigms

P300 wave is a Event Related Potential (ERP) elicited by external stimuli. Typically it is elicited accordingly the three different paradigms [67], as shown in figure 2.5.

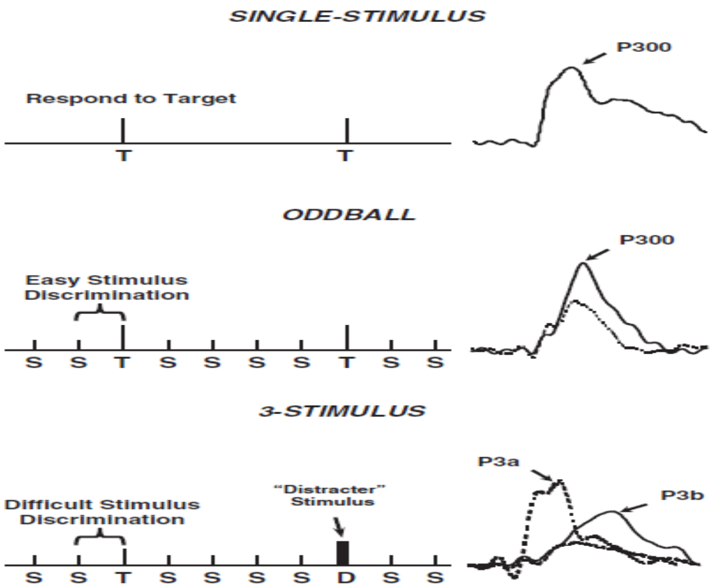


Figure 2.5: The three paradigms used to elicit the P300 ERP: Single Stimulus(upper), Oddball (center) e Three stimuli (lower).

First paradigm, known as the *Single Stimulus* paradigm, requires that the target is presented to the subject sporadically in absence of further stimuli. When the target is observed, a P300 potential can be determined.

The second protocol is the *Oddball Paradigm*. In this case, a sporadic target is displayed within a context of standard and repeated stimuli.

The third protocol, known as the *Three-Stimuli* paradigm, adds distracting stimuli to the oddball paradigm. The target stimulus is therefore rarely presented within a sequence of regular standard stimuli and sporadic distractor stimuli.

The P300 wave is measured by evaluating its amplitude and latency. Amplitude in the order of micro volt (μV) is defined as the difference between the average reference voltage before the stimulus and the highest positive peak of the ERP waveform within a time window (usually 250-500 ms). *Latency (ms)* is defined as the time elapsing from the start of the stimulus to the maximum positive amplitude point of the signal within a time window [68].

The P300 wave has an amplitude range between 6 and 20 μV and a latency between 250 and 400 ms. The peak signal persists at most for 100 ms.

According to studies by Emanuel Donchin [69], the P300 component indexes a brain activity when an incoming stimulus causes a change to a particular mental representation.

In particular, an attention driven comparison process evaluates the representation of a previous event in memory with the representation of the event just occurred. If no change in the stimulus is detected, the current mental model of the stimulus is kept in memory.

If a new stimulus is detected, a change or update of the stimulus representation is performed, resulting in presence of a positive peak of the P300 wave, as shown in figure 2.6.

2.4.7 The Oddball paradigm

Many works proved that a P300 based Brain Computer interface can be adopted for long time communication with people affected by latera amyotrophic syndrome. In particular it has been implemented an interface based on P300. (ALS) [70]. P300 wave was first described by Farwell and Donchin in 1988 [71], who defined the Oddball paradigm.

In the oddball paradigm a Bernoulli sequence is presented to the user. The

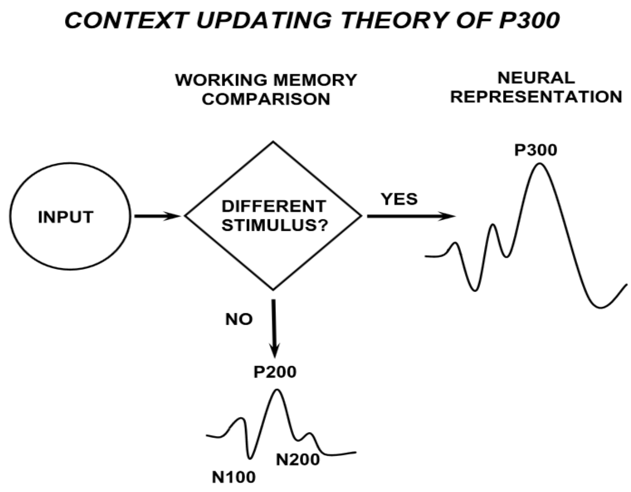


Figure 2.6: A schematic representation of a change or update of the stimulus representation in presence of a positive peak of the P300 wave

Bernouilli sequence is composed by a finite or infinite sequence of binary random variables, so it is a discrete-time stochastic process that takes only two values. The component variables X_i are identical and independent, , canonically 0 and 1.Those stimuli relate to two conditions, defined in this context as *frequent* and *infrequent*. User is asked to mentally count each time an infrequent condition is shown. [72]. To avoid habituation, conditions are presented in a random order. In Serby et Al. [73] conditions are presented as stimuli, represented as texture displayed over an interface organized as a matrix. Each element of the matrix represents an item. Stimuli cover the items element by element or in a row-column fashion. The use of matrix is particularly recommended to elicit a stronger P300 response [74].

The Classification methods

There are various classification algorithms for P300. The description of these methods assumes that the identification of a potential P300 evoked by an EEG can be considered a problem of binary classification with a discrimination function having a decision hyper-plane defined by the equation. 2.1.

$$wf(x) + b = 0 \tag{2.1}$$

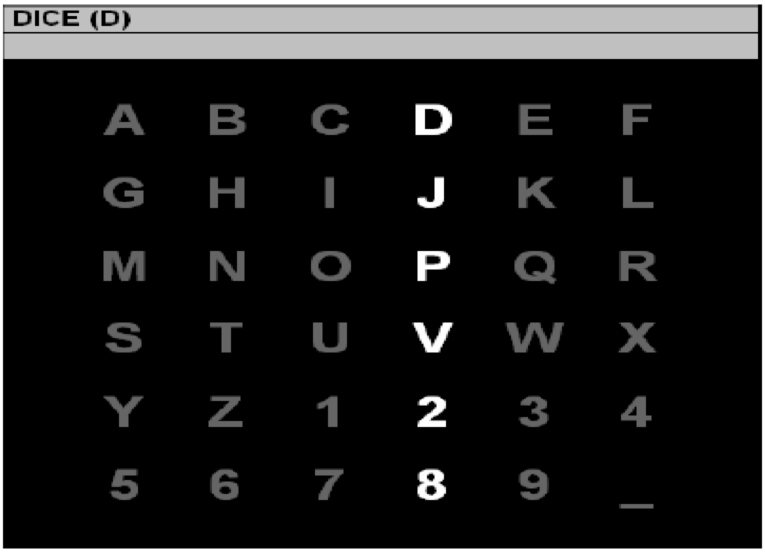


Figure 2.7: Farwell and Donchin Interface

where x is the *feature vector* to be classified, $f(\cdot)$ is the *projection function*, w is the *classification weight vector* and b is the *bias*. All methods uses different techniques to solve equation 2.1 with b and w. Some of the most used in literature are the *Fisher's Linear Discriminant Analysis* (FLDA); the *Stepwise Linear Discriminant Analysis* (SWLDA); the *Linear Support Vector Machine*; the *Feature Extraction* method; the *Artificial Neural Network*.

The *Fisher's Linear Discriminant Analysis* is one of the most widely used classification methods in BCI P300-based systems [75]. The basic idea is to determine the projection from an N-dimensional feature space to a one dimensional space for which the relationship between the variance between classes (*target* and *non-target* is maximized, while the intra-class variance is minimized. The optimal projection is defined by the2.2 equation.

$$w = (S_{-1} + S_1)^{-1}(\mu_{+1} - \mu_{-1}) \tag{2.2}$$

where S e μ define average and co-variance of classes to be divided.

The *Stepwise Linear Discriminant Analysis* can be defined as an extension of the FLDA with a feature selection filter. In particular, the method involves adding and removing terms from the linear discrimination model, thus producing an adaptable model to the data used for training [76].

The method starts adding the most statistically significant feature to a linear discrimination model. It follows a backward analysis to eliminate the least significant features. This process is repeated until the discrimination function includes a predetermined number of values.

The basic idea of the *Linear Support Vector Machine* is to determine the separation hyperplane, between two classes, to maximize the distance between the two classes [77]. By indicating the class labels with $y_i \in [\pm 1]$, equation 2.1 model can be reworded as follows:

$$y_i(wf(x_i) + b) + \eta_i \geq 1 \tag{2.3}$$

where $\eta_i > 0$ represent the separation margin between the two classes.

The *feature Extraction* based method is another possible linear [78] classifier. It consists of searching for a subspace that maximizes the *mutual information* between the set of projections

$Y = \left\{w^T f^i\right\}$and the corresponding T label set $t_i = \{-1, +1\}$.

As far as classification through a neural network is concerned, a multilevel artificial artificial neural network is considered forward with a single hidden layer and a sigmoidal activation functions [79], [80]. The classifier generated through such a neural network takes the form defined by the equation 2.4.

$$y(f, w, b) = \sum_{i=1}^{M} w_i^2 F\left(\sum_{j=1}^{N} w_{ji}^1 f_j + b_i\right) + b \tag{2.4}$$

where M is the number of hidden layer neurons, N is the number of observed features, b and w the threshold sets and coefficients weights.

2.4.8 Hybrid BCI

Hybrid BCIs are systems that interlace two BCI system or a BCI system with another system in sequence [81] or in parallel [82]. The design of a Hybrid BCI must respect the constraints that are related to the presentation of stimuli that evoke the particular brain responses to detect. In [83] researchers developed a hybrid interface based on gaze and motor imagery while [84] explored how the simultaneous gaze monitoring in conjunction with motor imagery could improve the overall performance of the system.

2.4.9 BCI and robotics

A Japanese study conducted by Tanaka collects the results of experiments to control an electric wheelchair through EEG signals [85].

Vora et al.[86] demonstrates the direct control of a robotic arm for everyday actions. Waytowich et al. [87] demonstrated that an EEG-based BCI provide accurate and reliable high-level control of a robotic manipulator and a human-machine interface between the human brain and the robotic manipulator is developed. [88] explores a control system for an automated robot, and authors developed a brain-actuated humanoid robot navigation system that uses an EEG-BCI to translate different human intentions into appropriate movement commands for robotic applications.

In [89], Chella et al. demonstrated the possibility to control a robot in a museum using BCI using a P300 interface.

Alimardani et al. [90] used the Geminoid robot as feedback for motor-imagery related tasks and [91] used an audio-visual immersive system to improve user's performance in a robot control task.

Robotic devices have also been used to provide control of robotic devices to facilitate stroke rehabilitation and control of orthoses [92], [93]. [92] describes a new non-invasive brain actuated wheelchair that relies on a P300 neuropsychological protocol and automated navigation.

2.5 The Biological Features

In this section I give an insight on the grounding theories to support the study described in chapters 4, 3 and 7.

The P300 component of the event-related potential is a waveform that can be extracted from the ongoing electroencephalogram through different paradigms [94] and it is associated with cognitive information processing (e.g. memory, attention, executive function) [95].

In statistics, the coefficient of determination, r^2 indicates the proportion of the variance in the dependent variable that is predictable from the independent variable [96]. In [45] it is used to evaluate sensorimotor cortices and in the mu- and beta-rhythm frequency bands during motor imagery. For Mark et Al.[97], r^2 is a good predictor of the P300 BCI performance for ALS people.

In a BCI context, the coefficient of determination is computed over signals that have been measured under two different task conditions. It is a measure of how well the original task condition ("user intent") may be inferred from a brain signal[3].

Neurological features have been used New hybrid BCIs explores the possibilities to calculate user's mental state during tasks and evaluating stress and fatigue during BCI tasks [98]. Stress, accordingly to [99], can be classified into: mental stress, emotional stress and physical stress.

The application of the described biological features, related with user's brain activity has been mainly explored during BCI task and not applied to the human robot interaction. My thesis, starting from the referred literature wants to explore how human robot interaction can be improved by the use of such features.

[100] introduced a Stress Index based on Shannon's Entropy to quantify the energy distribution from EEG Power Spectrum due to stressors. [101] used sample entropy to detect the difference in EEG signals due stress and fatigue condition.

[102] studied the correlation between eeg rhythms and additional information for characterizing user's mental states. In [103] a change in the pattern of alpha wave during drive fatigue has been reported. A method based on Shannon Entropy and Kullback-Leibler measured a relative quantification of fatigue based on spectral analysis of EEG [104].

Entropy describes the distribution of signal components and non-linear entropy methods describe the complexity of EEG signals [105].

[3]http://www.bci2000.org/wiki/index.php/Glossary - Accessed 15/05/2017.

2.6 Trust in robotics

In this section it is provided an insight on the concept of *trust*, to provide a theorical support to the study proposed in 7. The research into the meaning and the conditions of trust has spread from economics, psychology and philosophy to A.I. and robotics, in which it turns out to be fundamental. It spans the assessment of user's acceptance of technology and confidence in agents systems and e-services [106], [107] the control strategy in using artificial agents and robots [108] the design of human computer and robot interaction [109].

The research into trust provides such fields with interdisciplinary concepts and models, which are part and parcel of core questions centred on user experience and system design. Users trust in artificial systems is a factor of the likelihood of relying on the functions of computer aided decision systems, communication and transaction systems [110]. Trust calibration is crucial to select efficient and adaptive control strategies when users need choosing automation or supervision on the grounds of the reliability and the capability of the system [111].

As far as human robot interaction is concerned, the issue of trust is extended to the characteristics that make robots different from computer mediated or autonomous systems such as an embodiment that promotes in the user the sense of co-localization or even of co-presence with the artificial agent, the increased flexibility of high level functions supported by the sensory and the motor abilities of the robotic device, the widening of the operating environment in which more complex tasks are carried out, which come also close to those users deal with in ordinary experience [112].

To interact and cooperate with humans in their daily-life activities, robots should exhibit human-like "intelligence". This skill will substantially emerge from the interconnection of all the algorithms used to ensure cognitive and interaction capabilities [113]. Nevertheless, the initiative of a robot during a collaborative task with a human can influence the pace of interaction, the human response to attention cues, and the perceived engagement [114].

The issue of trust is scaled up in social robotics [115]. Robots are designed to embed or let the social intelligence abilities to arise in the situated interaction with the users, by which they can be recognized as team members or as companions, that is to say and partners on a par with human subjects, for work, assistance, and education [116].

Human-humanoid interaction is indeed a demanding test bed for the adverse effects on the interactions, which are due to the so-called uncanny valley [117]. For the same reason it is also a promising field to study the conditions at which the social human-robot interaction may be as much efficient as natural-like for the users.

There are in fact many definitions of trust, which however recognize the common character of trust in the willingness to bear the cost of relying upon another agent in uncertain or risky conditions at which the positive expectations on the latter weigh against one's vulnerability or the likelihood of failure [118].

The uncertainty is due to the lack of perfect knowledge of the intentions of other agents, and the risk is the perceived probability of a loss due to the opportunism and the free-riding of other agents who may defect instead of reciprocate, for the relation to be of mutual benefit, once having exploited some gain.

The application of the issue of trust in A.I. and Robotics emphasizes variously these elements of the concept of trust. Madsen and Gregor define trust as the confidence and will to act on the basis of the recommendations of an artificial intelligent decision aid. Lee and See construe trust as the procedural attitude that guides the reliance on automation to achieve one's goals at the boundary conditions, which limit the application of explicit rules and dedicated cognitive resources. The main limit of the referred state of the art consists in neglecting the use of neurological. features for trust assessment. My thesis wants to improve the current state of the art exploring a model of trust derived from neurological features.

[119] interprets trust as the willingness to accept the information and suggestions of a robot, and to be confident in its behaviour in high-risk situations.

2.7 Honest Signals

In this section it is provided the theorical grounding for the study proposed in Chapter 8.

The so-called social intelligence, as stated by Albrecht (2005), Ambady and

Rosenthal (1992) and Walker and Foley (1973) is manifested through the ability to exhibit and recognize social cues and behaviours. These two elements are an expression of attitudes which occur during social interactions through a multiplicity of nonverbal behaviours such as facial expressions and body postures as reported by Vinciarelli et al. (2009). The social signals, according to Richmond et al. (1991), represent, therefore, a continuous source of information for establishing emotions, moods, personality and other traits that are routinely used during the daily relationships.

Pentland and Heibeck (2008) refer to social signals as honest signals because they are a separate communication network that allows to accurately predict the nonverbal signs that typically occur during everyday human interactions. They also underline how some people often use signals as smile, frowns and fancy clothes as an expression of who they are or who they want to be. Many people use consciously this kind of nonverbal signals because they know the social effects that they can have on the communication.

[120] and [121] propose a set of four non-verbal movement features that characterize human behavior during the interaction.

[122] has extended the theory of honest signalling to strategic contexts in which humans are engaged in face-to-face or group interactions. Speed dating and salary negotiations are examples of face-to-face interactions, while tactical decision making and coalition membership shifting within and across groups are examples of social aggregates interactions when conflicting or competing interests hold.

Like in biology, honest signals are unconscious, in the sense that they do not involve conscious reasoning, normative or linguistic judgments, mandatory and costly in terms of cognitive resources.

[122] describes four types of honest signals by which agents tune, synchronize or change cognitive features that are socially salient like attention, understanding, interest, focus and openness. The first type collects signals of the *influence* that agents have in the interaction, which is displayed by the distribution of attention to control and orienteer the communication and the behaviour.

The second type collects *mimicry* signals that display the tuning of agents to each other, like nodding or leaning towards or away, providing feedback for mutual understanding and cooperation.

The third type collects signals of the *activity* by which agents maintain the interaction, which is displayed by the increase of the energy devoted to making gestures and sustaining the conversation.

The fourth type collects signals of the *consistency* of motivation and of determination of agents in pursuing some goals in the interaction. Those signals consist in modulating the energy applied to gestures and words and its distribution in time. To measure honest signals Pentland proposed a sociometric [123].

Discontinuous or smooth modulation and regular or irregular distribution are signal of the emphasis and the openness of agents as well as of the straightforwardness or conflicting interests of agents. Honest signals are traded back and forth by individuals face-to-face, within and across groups, hence they build "social circuits" or networks in which the trust and reliability needed by successful interaction are unconsciously settled.

On this account honest signalling serves as a machine whose function is drawing decisions and actions out of agents in such a way to solve coordination problems, when the information is not fully available to all of them.

Because signalling is mandatory [122] and [124] assumed that technological "socio-scopes" can track honest signals accurately and continuously as extraction of the of associated physical variables.

The described theory for honest signals have been mainly applied only in human-human interaction. In my thesis I intend to explore for the first time if this theory can be applied also in a human-human interaction mediated by a humanoid robot.

2.8 Emotions Representation

In this section it is provided a description of the theory used as basis for the study provided in Chapter 9.

Emotions can be described as the human response to adapt himself to the constant environmental stimuli deriving from one's circumstances, mood, or relationships with others. Emotions are the results of complex psychological and physiological systems activated by the perception of significant environmental events.

Many researches have been conducted in order to investigate how to represent

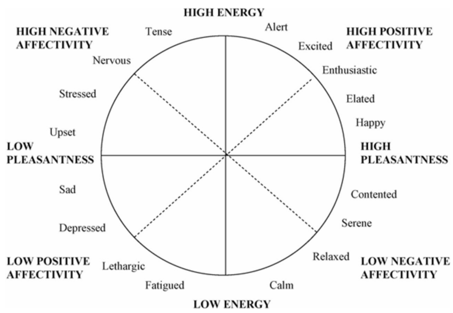


Figure 2.8: The Russel Circumplex. Each emotion can be maped on a bi-dimensional space based on arousal and valence.

emotions and how it is possible recognize them, as shown in [125] or [126]. Among the many approaches illustrated in psychology, the one based on emotional categories is one of the most often used in literature [127] , having focus on modelling emotions according to different and discrete emotion classes and labels. One of the most used approaches has been developed and described by Paul Ekman in [128].

In his work, Ekman claimed that all the emotions can be mapped into a six-dimensional space defined by six basic emotions, namely *anger*, *disgust*, *fear*, *joy*, *sadness* and *surprise*. Using this model it is possible to represent human emotions as a balance of the six basic emotions.

A two-dimensional, alternative to the Ekman's model, has been described by Russell et al. in [129], further extended by Posner et al. in [130]. In these works, authors defined a "circumflex model of affect" , aimed to accurately describe every possible emotion. In particular, authors suggest a two-dimensional structure in which an emotional event may vary quantitatively along two independent dimensional variables. One dimension (*valence*) describes the degree of pleasantness or unpleasantness, whilst the other dimension (*arousal*) indicates the intensity in terms of physiological activation. The combination of these two dimensions generates the subjective emotional feeling. By means of this model, every emotion can be defined as the combination of two different levels of arousal and valence. The emotional categories may be ordered along the aforementioned circumflex, as shown in Figure 2.8.

inthe study presetend in Chapter 9, the circumflex model has been used to describe the audience emotional responses, as well as the emotional content of the pieces of music of the composition. This allows us for discretizing the whole spectrum of emotions in four "areas" , one for each quadrant of the circumflex. Each of these areas was associated with a different color: from the first to the fourth quadrant, happiness (yellow), anger (red), sadness (blue) and serenity (green). These colors have been chosen for being associated with the correspondent emotions, accordingly to findings in [131] [132].

2.8.1 Music and robotics

There are several relevant works in the field of Robotics in musical applications. Among them, Hoffman et al. have defined the sense and outcomes of sharing an experience with a robot and this concept has been proposed as *Robotic Experience Companionship* (REC) [133].

Lim et al. proposed a unifying framework to generate emotions across voice, gesture, and music, by representing emotional states as a 4-parameter tuple of speed, intensity, regularity, and extent [134].

McCallum et al. stated that improvised musical interaction are able to provide an improvement of social presence and engagement during long term Human Robot Interaction (HRI) [135]. Burger et al. developed a simple robot with the goal to display its emotions by performing expressive movements during musical performances [136].

Brown et al. derived a framework for implementing happy and sad gestures on a humanoid robotic platform to measure user's predominant perceived emotions to enhances the social interaction [137]. This short list show the broad interest of the scientific community towards the use of (humanoid) robots for conveying emotions both from and to people in a musical context. The main limit of the presented literature consists in considering the interaction between a person and a robot, without taking into account the overall interaction of a group of people. My thesis wants to explore this neglected field of human robot interaction.

Chapter 3

The UnipaBCI Framework

In this chapter the unipaBCI framework is presented. It is a novel modular architecture for Brain Computer Interface. The design and implementation of this architecture has been a mandatory step to create a scalable, modular and reliable architecture to support researches described in the following chapters. A full description of the UnipaBCI Framework has been presented at the International Conference of Complex Intelligent and Software Intensive Systems [12]. The chapter is organized as follow: first the introduction on BCI framework is presented. Secondly, the UnipaBCI framework is described. Thirdly, UnipaBCI performance are evaluated in a laboratory setting.

3.1 Introduction

The development of BCI based application requires using framework with dedicated features and many BCI software both open source and proprietary have been developed nowadays.

BCI2000 has been developed since 2000. It is commonly used for data acquisition, and real time applications. BCI2000 supports different device for data acquisition systems, algorithms for signal processing and paradigms. It provides an SDK to develop new paradigms and module. The main workaround consists in developing new modules in a fast way since the framework become quite huge from its beginning.

OpenVibe [138] is an open-source software platform to design, test, and use brain-computer interfaces in real and virtual environments. Composed by many modules that can be connected and integrated to develop new BCI applications by source code or graphical block applications. Although is graphical approach based on module interconnections apears to be interesting, the interconnection with the g.usb amp, which was mainly used for my research, appears to be not stable and so it was not possible to use it as a software ground for my research.

IntendiX P300 is a proprietary software developed by g.tec medical engineering GmbH, Austria. It provides an EEG-based spelling software. It provides visual simulation, train the classifier and elaborate the data to extract a target character. It does not provide the opportunity to develop new modules and functions, since it is a property software.

All the proposed solutions appeared to have limitations for the purpose of my research and for this reason I started the development of a new BCI framework to support the subsequent researches.

3.2 Robotic application of BCIs

Brain-computer interfaces have been designed and developed to communicate and control peripherals that can replace or help users in movements and ations they may no longer be able to perform. A BCI system, in fact, allows to control an electronic or electromechanical device through a voluntary modulation of the user's brain activity, without the user's have to use his or her voluntary muscular apparatus. The interface known as Graz BCI, created by Pfurtscheller and colleagues in the early 2000's, uses instead the bands and sensory rhythms (SMR: Sensory Motor Rhythms) and their synchronization-desynchronization for some interesting applications. This interface has been used to control grasping actions.

This leaded to the design and implementation of UnipaBCI, a dedicated BCI framework, entirely designed and implemented to realize a real-time hybrid framework interfaceable with different hardware devices, User Interfaces, classification algorithms and robots. Such architecture has been mandatory to build up experiments and scenarios to test human robot interaction along different set.

3.3 The framework architecture

The UnipaBCI framework grounds on the architecture designed accordingly to the key points described in [139], where the mandatory features that a BCI framework must be equipped with are underlined. It is composed by five main modules: *Acquisition module*, *Signal processing*, *User Application*, *Device Controller* and *Control Application* as reported in figure 3.1. Each module communicates with the others through a TCP/IP network protocol.

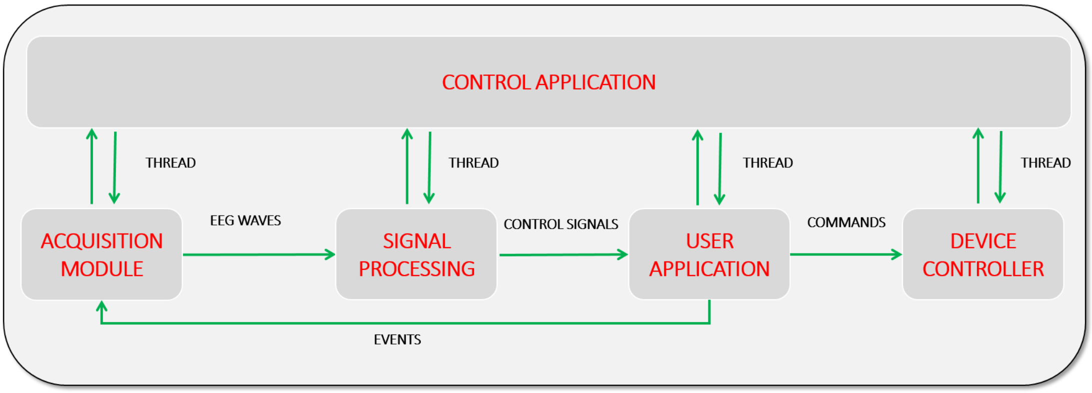


Figure 3.1: The architecture of the UnipaBCI.

The goal is to realize a real-time BCI system able to convert the brain signals in commands for the output device controlled. In particular, the User Application module setups the *BCI parameters* and provides a GUI for eliciting user's brain activity.

The brain signals, extracted in real-time from the user's brain, acquired by the *Acquisition* module, are elaborated and stored in dedicated multi-buffer structures to keep track of *BCI parameters*: channels, acquisition rates, filters and event markers.

The event markers are used by the control application to signal to all modules the beginning and the ending of the events related to the life cycle of the application.

The Signal Processing module extracts data from the buffers, elaborates them and sends the control signal obtained, through the User Application module, to the Device Control module, to manage the output device.

In current implementation, the gusbAmp BCI device produced by g.tec[1] has been used. The system has been designed to be modular to enable the development of new BCI paradigms and algorithms and to control different external devices.

3.4 The UnipaBCI framework

The system provides two operative modalities: *Training* to calibrate the system over the user and *Recognition* to use the system, after it is calibrated, to spell letters and symbols. Next paragraphs will provide an insight on the main parts of the architecture.

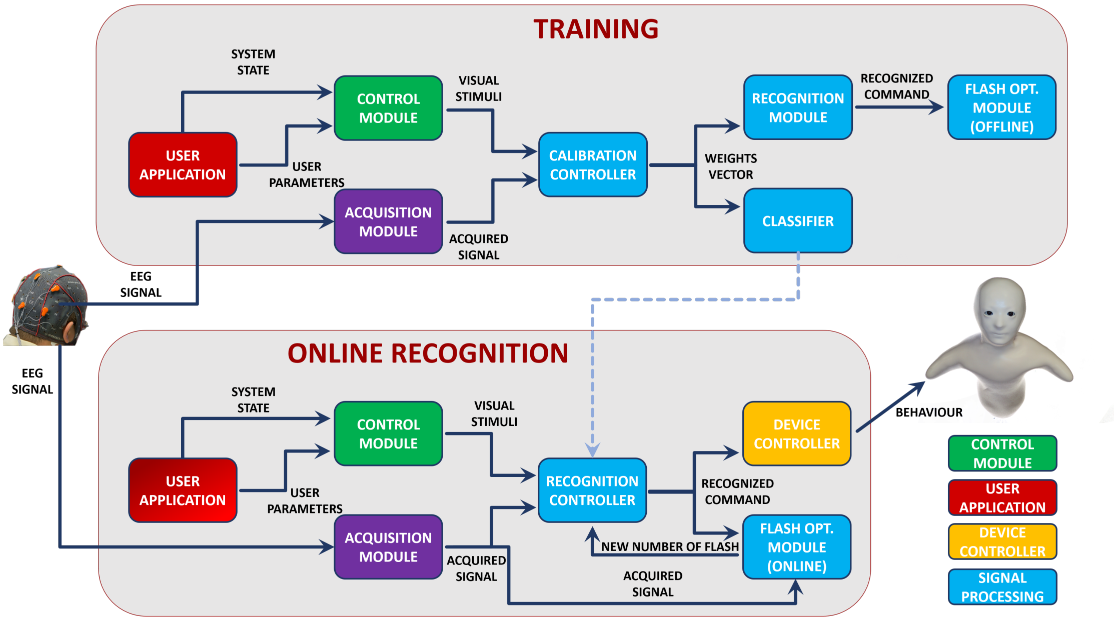


Figure 3.2: The architectural schema of UnipaBCI framework (*Each color of the blocks is related to a specific functionality*).

3.4.1 The Acquisition module

The system is compatible with the g.usbAmp, produced by g.tech, a high-performance and high-accuracy biosignal amplifier with 16 simultaneously sampled biosignal

[1]http://www.gtec.at/Products/Hardware-and-Accessories/g.USBampSpecs-features

channels with 24 bits, connected with USB socket to the PC (*for details please consult the Guger technology website (http://www.gtec.at)*. The maximum number of channels, as well as the sample rates and the applicable filters are fully dependent from the hardware. The framework theoretically works with more than one g.usbAmp at the same time but only single device tests have been done for the current implementation.

The Acquisition Module is used to connect to the G.usbAmp and to manage the recording of the electrical potentials, defined as raw signals, extracted in real-time from the brain of the user. The acquisition rate was set at 256 Hz and it has been implemented a Notch filter of 50 Hz. Raw signals are band-passed between 1 and 50 Hz with a $6 - th$ Order Butterworth filter.

The g.usbAmp supports the following sampling rates: 32, 64, 128, 256, 512, 600, 1200, 2400, 4800, 9600, 19200, and 38400 Hz. The Notch Filtering could be applied in the ranges of 48-52 Hz and 58-62 Hz. The bandpass filters can be applied between 0 and 2000Hz.

3.4.2 The Signal Processing module

The signal Processing Module is used to extract relevant features from the raw signal. This module is composed by a pre-processing module, used to remove artifacts from the signal, a training module, dedicated to the computation of weights of the classifier and a recognition module used to discriminate user's brain activity. In figure 3.3 is described the architectural schema of the Signal Processing Module.

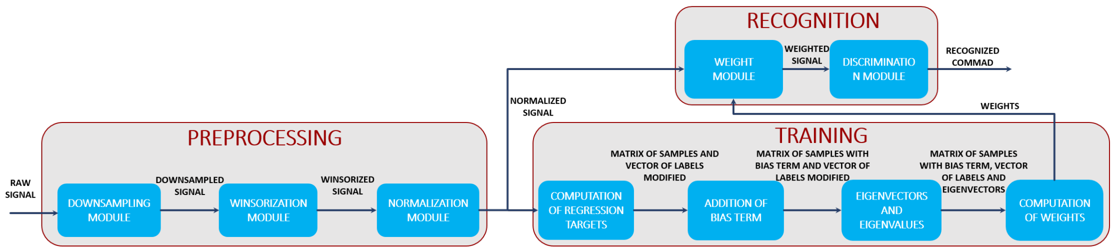


Figure 3.3: The architectural schema of Signal Processing Module.

Signals, acquired and digitalized by the Acquisition module are down-sampled to reduce data dimensionality. To avoid aliasing effects, a preliminary study to find the best reduction factor has been carried out and, based on findings in [140] and [141], a reduction factor $n = 4$ has been set up. In this way, signals originally acquired at 256Hz are stored with a sample frequency of 64Hz.

The Windsor filter is applied to the signal to cut off the outliers values reducing their bias effect. *The Winsorization* module defines a threshold that considers values as outliers, based on median and inter-quartile range and it takes into account only 1% of these values. The signal is then normalized using a z-score method [142]. The algorithm works by subtracting from each feature the mean μ and dividing it for the standard deviation σ.

The Training Module implements a Bayesian Linear Classifier [143] that outputs a weights vector to classify the normalized samples obtained from the Preprocessing Module.

The brain signals acquired are synchronized with the Visual Stimuli related to two class labels: *target* and *non-target*. The brain signals and the class labels are used by the classifier to discriminate the membership for each stimulus.

Brain signals are inserted in a structure called BSP (Brain Save Potentials), with dimension [F * N] where F represents the features extracted and N is the number of normalized samples extracted for each epoch. The epoch size has been set to 600 ms.

The *Regression Targets Computation* calculates the regressions targets from the normalized samples using the Bayesian Linear Discriminant Analysis (BLDA) [144] .

BLDA is based on the assumption that regression targets are set as $\frac{N}{N_i}$ where N is the total training samples and N_i is the total number of samples for the $i-th$class. The *Bias Terms addiction* consists in the introduction of a bias term, as a systemic error due to data modelling.The *eigenvectors and eigenvalues calculation* calculates the eigenvectors and the eigenvalues for $X' = X \cdot X^T$

The weights are calculated using the following equation:

$$p(D|\beta, w) = \left(\frac{\beta}{2\pi}\right)^{\frac{N}{2}} exp(-\frac{\beta}{2}||X^T w - t||^2) \tag{3.1}$$

where **t** represent the *regression targets*, **X** are the in the training matrix, **D** is the pair$(\mathbf{X}, \mathbf{t})$, β is the inverse of the additive noise and $\mathbf{N}$ is the total number of samples.

To calculate the posterior distribution w it is necessary to calculate the conditional expectation for w given α. This distribution is expressed as a multivariate Gaussian distribution with zero mean, expressed as:

$$p\left(w|\alpha\right) = \prod_{i=1}^{D} \left(\frac{\alpha_i}{2\pi}\right)^{\frac{1}{2}} exp\left(-\frac{1}{2}w^T I'(\alpha)w\right) \tag{3.2}$$

where $\mathbf{I}'\left(\alpha\right)$ is a diagonal matrix with dimension D x D; α_i is the inverse of the variance of the distribution of each weight w_i and it's used to estimate the importance of w_i in the regression model[145].

From this equations, the algorithm iterates the process and calculates the α and β values until they are lower then an error ϵ. *The Recognition* module is used to recognize the command selected by the user with the BCI device after w is calculated.

The recognition module determines a new vector of signal samples as:

$$m_k = \sum w * x. \tag{3.3}$$

where w is the weight vector and x is the original brain signal. For each $k-th$ symbol of the interface the x acquired during the presentation of the $k-th$ stimulus are summed with the corresponding m_k:

$$v_k = m_k + x_k \tag{3.4}$$

The k-th symbol, corresponding to the maximum v_k is selected as output of the Signal Processing module.

3.4.3 The User Application module

The User Application module, shown in figure 3.4, consists of the Operator Console and the User Interface. The Operator Console is used by the experimenter to set system parameters, choose operational modality (*training*, *recognition*), monitor the experiment and set the network address of external devices.

The User Interface provides the stimuli interface and it consists by a symbolic or alphabetical matrix whose dimension and icons are dynamically selected by the user.

The user aplication module supports single stimulus (here defined as single square modality) and oddball(here defined as row-column modality) elicitations. Each item is highlighted with an icon showing the Einstein face, to increase ERPs response. In fact, findings in [146] demonstrated that well known faces strengthen the P300 response.

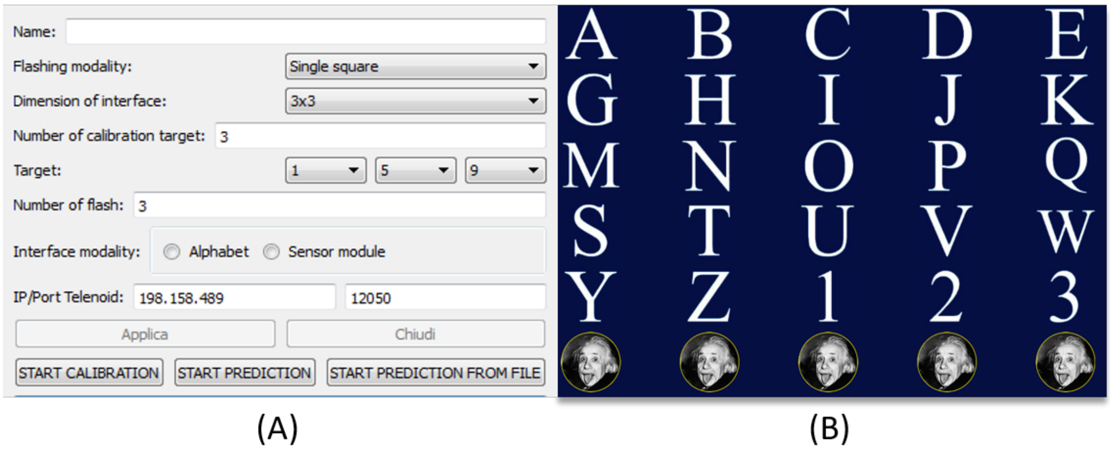


Figure 3.4: The Operator Console (A) and the User Interface (B).

3.4.4 The Control Application

The *Control application* module, shown in figure 3.5 consists of the *States Manager* as the main controller and by three sub-controllers *the Acquisition Controller*, the *Interface Controller* and the *Device Controller*.

The *States Manager* is the core of the whole platform. It is the main thread, managing all system states. After the platform has been started, it could be in *Training State* or *Recognition State*. For each of the states it manages the execution and synchronization of the sub-controller.

In *Training State* it will activate the *Acquisition Controller* and the *Interface Controller*.

In *Recognition State* it will activate the *Acquisition Controller*, the *Interface Controller* and the *Device Controller*. The *Acquisition Controller* is used to manage the EEG signal recording and storing. The *Interface Controller* manages the execution of the User Interface and the stimuli synchronization. The *Device Controller* controls the external device, used as final output of the system.

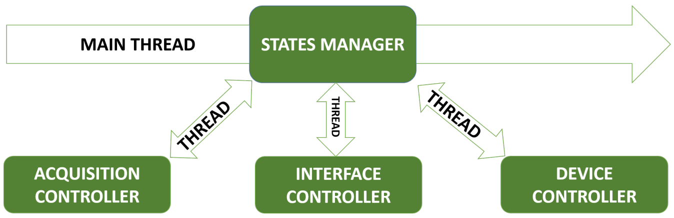


Figure 3.5: The Control Application module.

3.4.5 The Flash Optimization module

Accordingly to personal differences (e.g. *BCI skills*, *level of concentration*, *stress*) it is not possible to have a prior definition of the right number of stimuli needed to discriminate the P300 potentials.

For this reason, in order to find the correct number of stimuli (flashes) to be presented to the user during the trials, two modules have been implemented: one for the training session and the other one for the recognition session.

The Offline Flash Optimization module is used to identify the best number of stimuli to be presented to the user based on the percentage of symbols correctly classified in the training set. This modules merges the vector of weights w with the target id representing the expected symbol of the interface user must mentally select.

The optimization module calculates the minimum number of flashes needed for the best precision of the system, intended as the probability of correctly classify the id of the target.

The Online Flash Optimization is used during the Recognition phase, when there are no prior informations on the target id that user is going to select.

For the target id selected the Online Flash Optimization evaluates the minimum number of stimuli for correct selection and set the next iteration stimuli to that value.

3.4.6 The Device Controller module

The Device Controller module is used to send commands to an external device connected to the system. In this study, the Device Controller Module sends commands to the Telenoid R1 robot used as an embodied communicator.

To detach the platform from the device connected it has been implemented a network interface based on TCP/IP for connection. In this way the robot doesn't need to be in the same physical place as user. Each command selected, it is sent to an *Accumulation Buffer*, until a termination character is selected. The obtained string is translated in an audio file using *Text-To-Speech* API.

The Device module also initializes the Network Connection and sends the vocal file to the robot to be reproduced. At the end of the reproduction, Telenoid sends a termination ack to the device module which informs the device controller of the correct termination of the behaviour and enables the system to perform further operations.

3.5 Online Evaluation and validation

The experiment has been conducted in a controlled light room under the supervision of an experimenter. The experimenter controlled all the phases of the experiment and instructed the participants.

Participants were instructed to avoid movements, blinking, swallowing and speaking. To limit the artifacts, it was asked them to focus on the GUI and to mentally count each time the expected item was highlighted.

Participants wear a g.Cap with electrodes in Cz, Pz, C3, C4, accordingly the 10-20 montage system [147]. This montage has been empirically eliged because it provided best results, even with a small number of electrodes.

Participants sit at a distance of 40 cm from the interface monitor, a Telenoid robot was connected to the system to provide feedback to the participants on the ongoing of the experiment. In figure 3.6is reported a schematic description of the experimental environment setup. The experiment has been conducted on 20

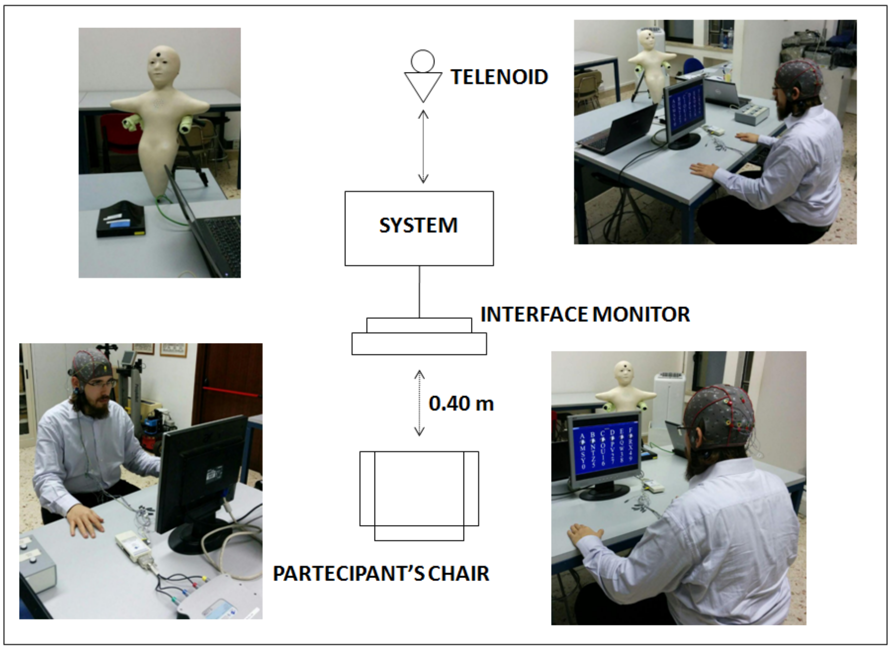


Figure 3.6: The experimental environment setup.

healthy participants with a median age of 23 years, 15 male and 5 female,with an average scholarization of 14 years. 10 of them have been defined as *Expert* since they had a prior knowledge in using BCI (*more then 10 hours in controlling BCI*) while the others five had no prior knowledge of BCI systems. All the participants were informed of the purposes of the study and gave their written informed consent.

3.5.1 The precision of the system

An experimental session with expert and non expert BCI users has been conducted. In this experiment it has been investigated the precision of the UnipaBCI framework accordingly to personal skills and using different operative modality: *single square* and *row-column* with a 5x5 alphabetic interface.

In single square mode each target was highlighted separately while in row-column mode for each target, all the row or all the column containing it were highlighted.

The first step consisted in the calibration to train the system over users. For this step a 5x5 matrix formed by the first 20 letters of the English alphabet was

Row - Column mode		S1	S2	S3	S4	S5	S6	S7	S8	S9
Experts Precision	80.00%	80.00%	100.00%	90.00%	80.00%	100.00%	100.00%	80.00%	100.00%	90.00%
Beginners Precision	90.00%	80.00%	70.00%	90.00%	100.00%	80.00%	100.00%	70.00%	80.00%	90.00%
Experts avg flashes	5	6	6	5	5	4	4	5	3	7
Beginners avg flashes	9	8	7	9	10	8	10	7	8	9

Table 3.1: Results for expert and beginners in terms of precision and average number of flash with a 5x5 interface and row-column modality.

used. The calibration session was formed by two trials, and each trial was formed by a sequence of three items. A total of 10 enlightenment for each item have been shown to each subject. At the end of the calibration session, the system calculated the attended percentage of success to be used during the recognition session to achieve a prediction of 100% of success for classification.

All participants successful completed the calibration phase with success. This result suggests the BLDA algorithm was able to correctly classify the provided s.

In the recognition modality, participants were asked to recognize a total of 20 items randomly chosen by the experimenter from a 5x5 alphabetic interface. 10 item were presented in *row-column modality* and 10 in *single square mode.*

The stimuli presentation modality between users has been randomized. Each item was highlighted from 1 to 10 times for 125 ms accordingly the results provided by the Optimization Module. Between two subsequently stimuli, a random time interval between 175 ms and 225 ms was set to avoid users' habituation.

Table 3.1 reports results obtained in row-column (RC) alphabetic modality. Experts achieved an hit rate of 90.00%$\pm$ 9.43% while beginners successfully completed the task with a precision of 85.00%$\pm$ 10.80%.

Accordingly to the row-column paradigm used, for a 5X5 interface, the total number of possible target is 10. The average number of flash, set by the Optimization module was 5$\pm$ 1 stimuli for Expert and 9$\pm$ 1 stimuli for Beginners (all values are averaged to the closest integer by the Optimization module). In table 3.2 are reported results for the single square modality. Experts achieved an average precision of 85.00% $\pm$ 13.54% and 80.00% $\pm$ 9.43% for non expert users.

The average number of flash, set by the Optimization module was 5 ± 1 stimuli for Expert and7 ± 1 stimuli for Beginners.

A statistical analysis of results provided in table 3.1 and 3.2 are reported in section 8.3.

Single Square mode	S1	S2	S3	S4	S5	S6	S7	S8	S9
Experts Precision	100.00%	90.00%	90.00%	80.00%	100.00%	80.00%	80.00%	70.00%	60.00%	100.00%
Beginners Precision	80.00%	80.00%	70.00%	80.00%	70.00%	80.00%	80.00%	70.00%	90.00%	100.00%
Experts avg flashes	6	4	5	7	5	4	4	5	6	7
Beginnersa vg flashes	4	6	7	8	9	8	6	8	7	7

Table 3.2: Results for expert and beginners in terms of precision and average number of flash with a 5x5 interface and Single Square modality

3.5.2 The *speedup factor*

The speedup factor represents a measurement of the system performances in terms of bit-rates defined as the number of correctly spelled items. The Optimization Module provides adaptation on the number of stimuli accordingly to the classifier results. in particular, the number of stimuli is set at the beginning of each session and whenever the classifier prediction reaches 100% the correspond symbol is provided as output and the remaining stimuli are not presented to the user.

The **speedup factor** F_s represents the differences in speed with and without the optimization module active and improves the bit-rates of UnipaBCI. It could be described as:

$$F_S = \frac{S * T_{Opt} * St * I}{S * T_{NoOpt} * St * I}$$

Where S is the number of participants, T_{Opt} is the total number of target with Optimization Module active, T_{NoOpt} is the total number of target with Optimization Module inactive St is the stimuli duration, I is the inter-stimuli duration.

Each stimulus lasted $125ms$, the inter-stimulus was set between 175 and 225 ms and a total of 10 users accomplished a total of 200 trials for each modality (Row - Column and Single Square).

Without the Optimization module active, the number of flash was set to 10, while in table 3.1 and 3.2 are reported the average number of flash used during the experiment In consideration of this, the speedup factor with the optimization module active was 1.49 in Row-Column mode and 1.62 in Single Square mode.

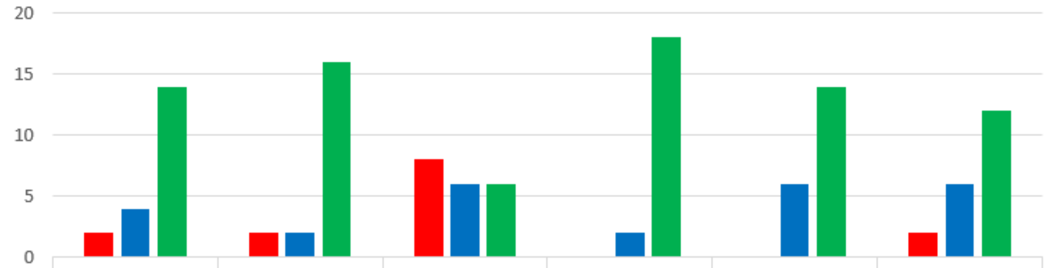
3.5.3 Questionnaire analysis

At the end of the experiment, it has been asked have asked participants to fill a questionnaire consisting of 6 questions to be answered with 3-points Likert scales.

All 20 participants filled the questionnaire.

The purpose of the questionnaire was to examine the user acceptance of UnipaBCI from a qualitative point of view and to have the confirmation that the proposed system meets all the necessary requirement for a BCI framework.

From results shown in figure 3.7, it is clearly possible to deduce that participants find unipaBCI easy to use (Q1,Q2); the experiment was well prepared (Q4, Q5). Although participants found quite demanding to focus during the experiments (Q3), 12/20 participants were overall satisfied while only two participants were not satisfied (Q6).



	Q1:System is easy to use?	Q2: Beginners can use the system?	Q3: Using UnipaBCI do not require lots of focus?	Q4: Experimental Environment is well lighted?	Q5: Flash Timing is proper?	Q6: I am overall satisfied by UnipaBCI?
I Disagree	2	2	8	0	0	2
Neutral	4	2	6	2	6	6
I Agree	14	16	6	18	14	12

Figure 3.7: Results of the questionnaire filled from participants after the experiments.

3.5.4 Real Time Capabilities of the UnipaBCI

The framework has been developed using Microsoft Visual Studio 2012 to provide a complete IDE for the code writing. The operating system used to implement UnipaBCI is Windows 10 since it provides more support to the API of the g.usbAmp hardware device used to test the framework.

C++ was selected as language, because it introduces a low performance overhead at run-time. GUI and other visual components are implemented in OpenGL. To test the real time capabilities for the UnipaBCI framework it has been followed the approach stated in [148]. Test was run on a Intel Core i5-4200M wit 8Gb at 2.5Ghz, with 8GB ram and integrated graphic card. The processor load was always $< 30\%$.

Our test has been run at 256Hz so the acquisition module received a new sample every 39,06 ms. The ADC latency, expressed as the time to acquire a signal for each of 16 channel, was below 3.17ms. This result proves that the system is able to process every sample and store it before a new one is acquire.

The processing latency, expressed as time to get a prediction from the BLDA algorithm was 4.022$\pm$ 0.59s. The output latency expressed as the time to send the command to the robot was < 22ms.

These results suggest that the most demanding operation is represented by the BLDA algorithm while others operations are done without any significant jitter so UnipaBCI framework could be effectively used for real-time operations since prediction is done at the end of each trial.

3.6 Discussion

An unpaired t-test with $\alpha = 0.01$ has been conducted to asses the obtained results. Accordingly to our findings the precision of the system is not statistically affected (t(40) = 1.7471 p= 0.0887) by flashing modality (Row Columns vs Single Square) so it is possible to conclude the system could be used indifferently in both modalities even if users obtained best results in Row Column modality.

Both expert and beginner users achieved good results using UnipaBCI and expertness doesn't seems to statistically affect the performance both in row-square modality (t(20) = 1.1028 p = 0.2846) and in single square modality (t(20)=0.9583 p = 0.3506). It is clearly possible to conclude that the system is performing accurately and that it works even with non expert participants.

Accordingly to the questionnaire, all users, even beginners, found the system easy to use. They noticed quite difficult to follow the flash speeds but they were overall satisfied by UnipaBCI.

Accordingly to the performance test, the system could be used in real-time and present good signal acquisition rate and fast connection to the external agent. The most demanding operation is represented by the features classification.

In conclusion the UnipaBCI has demonstrated to be a powerful and reliable system able to work efficiently with visual evoked potentials showing a low error rate with Expert and Non Expert participants. Moreover only a low number of electrodes have been used.

The flash optimization module proved that the system was able to adapt over the user expertness by finding the optimal number of stimuli to provide for each trial to avoid stress in the user.

The system proved to be able to work in real-time and suitable to be the ground to develop new human robot interaction studies based on Brain Computer Interface in a fast and reliable way.

Part II

Robot as Avatar

Chapter 4

Evaluating the Biofeedback factors in a human humanoid robot interaction

In this chapter is presented a study conducted with ALS participants to evaluate the interaction with a NAO robot using Brain Computer Interface for accomplishing their needs. The study presented in this chapter declines the main research question, proposed in chapter 1 in two sub-questions:

1. *Is the NAO Robot able to help a human in accomplishing his needs?*
2. *Is it possible to derive the humans' mental state during the interaction with a robot?*

The first sub-question investigates the acceptance of a BCI based interaction to control a humanoid robot for fulfilling ALS's needs, comparing these results with an healthy control group.

The second sub-question has been evaluated defining the *biofeedback factor* B_f, representing the users' mental state during the interaction with the humanoid robot. This factor based on users' Attention, Intention and Focus as feedback of users' engagement during the interaction with a robot.

To exclude cognitive disorders and assess motivation the neuropsichiatrists of the ALS Center of Palermo performed a test battery before the experiment. The experiment and the test battery were approved by "the Institutional Palermo-1 Ethics Review Board, Ospedale Policlinico Giaccone, Palermo Italy".

After a preliminary interview to understand patients' needs, it has been chosen as experimental task an interaction where the robot had to reach and grasp a glass of water. The action was triggered by the BCI selection of the command corresponding to the correct action. During the execution of the tasks, I acquire the Intention, the Intention and focus which represented a function of the user's Biological feedback.

This study has been conducted in collaboration with the department of Experimental Biomedicine and Neuroscience, University of Palermo (Italy) with ALS locked in patients. The case of study described in this chapter has been published in the International journal Frontiers in Human Neuroscience [149] and in the IEEE Transactions on Neural Systems and Rehabilitation Engineering [150].

The chapter is organized as follow: first the architecture design is provided. Secondly, the motivation and neurological assessment is provided. Finally, the experimental protocol and the results are described.

4.1 The Architecture design

The system has been implemented as an istance of the UnipaBCI framework, the general purpose architecture described in chapter 3. To explore the role of biological feedback in human-robot interaction it has been created a use case scenario where a user had to control a humanoid robot to reach and grasp a glass of water using high-level BCI commands.

The robot used was the NAO Robot, an autonomous, programmable humanoid robot developed by Aldebaran Robotics[1].

The system was also equipped with a Tobii Eyex^{TM} [2], a device ables to track in real time the movement of users'gazes at a frequency of 50 hertz.

The UnipaBCI default Application module has been modified defining a new User Interface formed by a 3x3 matrix, as shown in figure 4.1, where each item

[1] https://www.ald.softbankrobotics.com/en/robots/nao
[2] https://www.tobii.com/tech/products/platforms/

is a command for the robot. The interface implemented two different types of behaviours corresponding to robot actions:

- *Low Level Behaviours*: to control the robot by moving it in different directions;
- *High Level Behaviours*: to grasp and give back an item.

My preliminary tests demonstrated that a blue background screen with white symbolic icons provided the right mixture of concentration and contrast to elicit tasks. During trials, each icons is randomly replaced by Einstein's face. Two new modules have been designed and implemented:

- A Biofeedback System based on neurological states and gazes position;
- A robotic system that translates user's mental activity into behaviours for the humanoid robot.

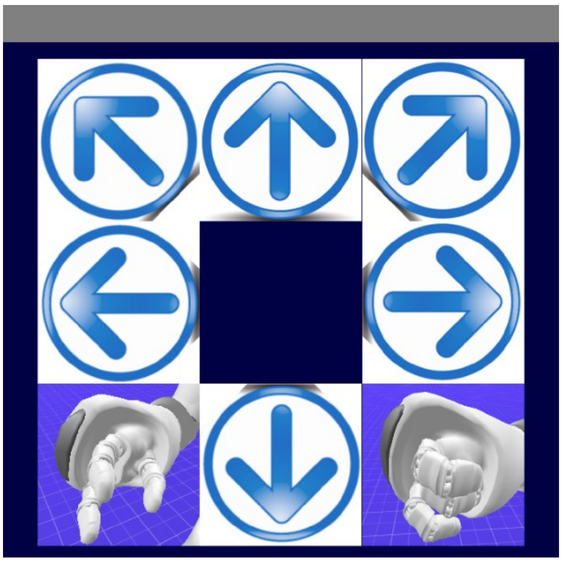


Figure 4.1: The VEP User Interface with 6 Low Level Robot Behaviours: *Left (A1), Right(A3), Up (A2), Down(C2), Turn Left (B1), Turn Right(B3)*) and 2 High Level Robot Behaviours (*Grasp* (C1), *Give* (C3).

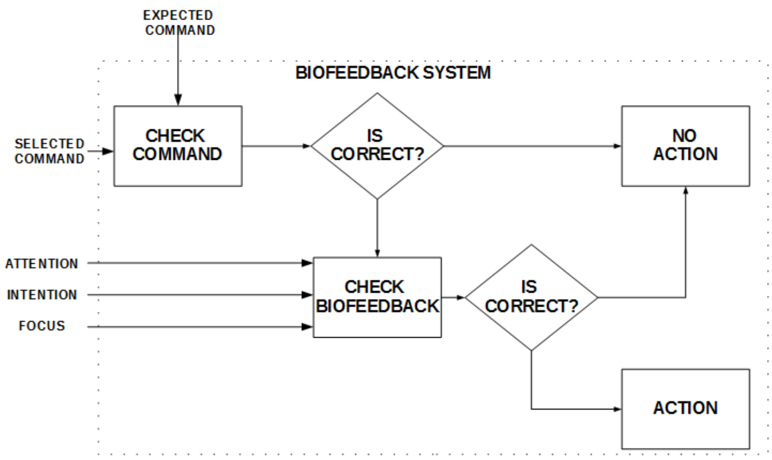


Figure 4.2: The flow chart for execution of Biofeedback System .

4.1.1 The Biofeedback System

The Biofeedback System comprises 4 biological parameters based on the user's eyes and brain activity: *Mental Attention*, *Intention*, *Visual Focus* and *Stress*. It is composed of modules able to extract these parameters from the user at run-time. In figure 4.2 the whole Biofeedback System is described.

Accordingly to the paradigm chosen (see section 4.4), users were asked to select a prior known commands C_e, defined by experimenters.

The Biofeedback System checks the equivalence between commands C and C_e and execute an action only if B_f is greater than a fixed threshold of 60%. Otherwise no action is executed by the robot.

The threshold has been fixed by a process of trial and error. The Biofeedback System merges the different bio-parameters, as a weighted sum, and gives as output B_f a value in percentage between 0% and 100% as function of the user's inner state as reported in equation 4.1.1:

$$B_f = \begin{cases} \alpha_1 A + \alpha_2 F + \alpha_3 I & if(B_f > 60\%) \quad \&\& \quad (C_e == C) \\ & 0 \quad otherwise \end{cases} \tag{4.1}$$

where A is the level of Attention, I is the intention, F is the Focus and C is the selected command through the User Interface. α_1 is the weight for Attention, with value $\frac{2}{5}$, α_2 is the weight value for Visual Focus, with value $\frac{2}{5}$ and α_3 is the weight value for Intention with value $\frac{1}{5}$.

Weights for visual focus, user's Attention and Intention have been set accordingly to findings in preliminary tests. The following paragraphs describe each element of equation (2).

Measuring Attention The system provides 9 possible commands, so a 9 classes decomposition is needed for a correct classification.

In this case, an analysis derived from Fisher's Linear Discriminant is used to obtain a subspace which contains all of the class variability with a technique known as *One Versus Rest* [151], which maximizes the inter class variance by measuring the distance of one class from all the others.

The techniques above can determine the features in terms of timing and position over the brain.

User's Attention A (expressed as percentage), is classified during the current task as the sum of recognized stimuli est_i for symbol s_i, where s_i represents the $i-th$ symbol of the interface:

$$A = max(\sum_{i=0}^{N} est_i) \tag{4.2}$$

Measuring Intention The *Intention* value **I** is measured from the correlation factor, r^2, of the P300 wave during ERPs. This value varies both within and across participants on different trials. To normalize value, the $I_{min} = min(r^2)$ and $I_{max} = max(r^2)$ are calculated during calibration phase.

The Intention I_t (*expressed in percentage*) represents, for each trial, the Intention of the user in accomplishing the task. It is calculated as described in equation 4.3:

$$I_t = 100 \cdot \frac{r_t^2 - min(I)}{max(I) - min(I)} \tag{4.3}$$

where r_t^2 is the current value of the correlation factor. The new values I_{min} and I_{max} are calculated as $I_{min} = min(I_{min}, I_t)$ and $I_{max} = max(I_{max}, I_t)$.

Measuring visual Focus An eye tracker that follows the user's gazes has been used to evaluate visual focus. For each trial the visual Focus Module is able to detect if the user is focusing on the expected part of the interface or not.

The area is divided in three sub-areas. The first one is the outer focus, representing the points external to the boundary box containing the symbol of the interface. The second area is the lateral focus area, covering the external part of the symbol and representing a partial focusing of the user.

The third area represents the central focus and represents a complete focus from the user. The gaze points have been clustered according to the three areas using 3 classes k-means.

The visual focus V, extracted from user's gaze is evaluated according to the following equation:

$$V = \alpha_1 F_c + \alpha_2 F_l + \alpha_3 F_o \quad with \quad V \epsilon [0 - 100\%] \tag{4.4}$$

where F_c represent the percentage of central focus, F_l represents the the percentage of lateral focus and F_o represents the percentage of outer focus. $\alpha_1 = 1$, $\alpha_2 = 1/2$, $\alpha_3 = -1$ are three weights to modulate the output of the Focus Module accordingly to the areas focused. Since it is expected that user focus on icon central area, the point inside the central area are considered as full focus and weighted 1, the point in the later area are considered as partial focus and weighted as an half while the point outside the expected icon are considered as distraction and therefore subtracted.

Measuring Entropy Accordingly to literature findings [152], it is possible to find relationship between between entropy and the complexity of brain signal and a decrease in entropy leads to an increasing in the amount of brain signal information. on the other side, high entropy appears to be related to stressful situations [101].

Starting from these assumptions, I conducted a post-hoc analysis on sample entropy to measure users' stress during the experiment. To obtain sample entropy,

Signals acquired separately from each electrode were averaged into one signal as mean of the original signals using a technique known as "*Grand mean*" [153] to obtain a single channel signal:

$$\bar{X} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2 + ... + n_k\bar{x}_k}{n_1 + n_2 + ... + n_k} \tag{4.5}$$

The obtained signal is decomposed into its constituent frequencies using the *Daubechies* D4 Wavelet Transform to extract the brain rhythms as:

$$H(X)_{rythm} = -c\sum_{i=0}^{m} p(x_i)ln(p(x_i)) \tag{4.6}$$

where c is a positive constant, x_i is a 256Hz sample of brain signal, $p(x_i)$ is the probability of $x_i \epsilon X$ with X defined as a set of finite random variables. x_i, moreover, must satisfy the following equation:

$$\sum_{i=0}^{m} p(x_i) = 1 \tag{4.7}$$

The normalized values of entropy were averaged over the constituent bands to reconstruct the overall activity over all the rhythms. The entropy value for the $i - th$ interaction is evaluated as:

$$S_j = \frac{\sum_{i=j-k}^{j} S_i}{k} with j > k \tag{4.8}$$

Where the entropy level for the $j - th$ interaction is represented as the entropy level of the past k interaction.

4.1.2 The Robotic System

The robotic system represents the output of the system. In particular, two interchangeable control modes been developed:

- Navigation mode where the NAO walks in 6 directions (forward, backward, turn left, turn right, rotate right and rotate left), controlled by user's mental activity;
- High-Control mode where the robot can accomplish complex behaviours autonomously such as grasping an object, locating the user, and bringing the

object back to the user.

Navigation Mode In Navigation mode, the robot acts as an avatar of the user, who can discover the environment through the webcam present in robot's head. Thanks to the network architecture provided, the vision system can be used remotely. For safety reasons, it has been implemented a reactive system that incorporates sensors such as sonars and bumpers called reactive system.

The reactive system controls the navigation behaviors to assure that the robot always acts safely, avoiding obstacles or stopping if a danger condition is detected.

This operation modality has been not used in the experiment described in section 4.4 and it is only described for the sake of completeness.

High-Control mode The high control mode is able to interpret high level commands received from BCI triggering autonomous behaviours on NAO to accomplish the corresponding task. Nao can execute two different tasks:

- Grasping an object;
- Giving an object.

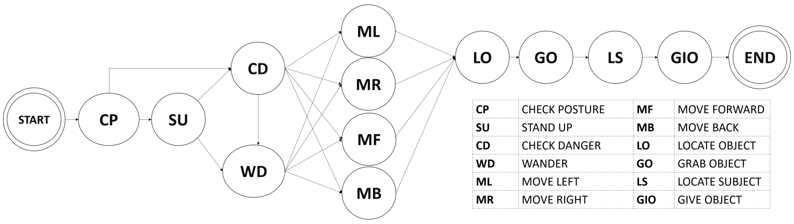
The high-level commands entail a combination of simple behaviours and robot-induced activities such as obstacle detection and avoidance. Before moving, the robot activates its sensors to evaluate obstacles or potentially dangerous situations, represented by $O = \frac{Revealed\_Obstacles}{distance}$.

If no obstacles are detected, this function returns zero. Otherwise, it will return a value which is inversely proportional to the distance. In case $O < T_{obstacle}$, where $T_{obstacle}$ is the minimum distance to consider a reaction safe, the S_k command is activated on the robot if the B_f biological factor is not null. Summing up, the command R_k executed by robot is function of:

$$R_k = f(S_k, B_f, O) \tag{4.9}$$

where S_k represents the $k - th$ command associated with the stimulus and B_f is the factor based on the user's neurological and biological states recorded by the Biofeedback System.

The robot decomposes the complex behaviours in many sub-operations executed independently in a state machine fashion, as shown in figure 4.3.



CP	CHECK POSTURE	MF	MOVE FORWARD
SU	STAND UP	MB	MOVE BACK
CD	CHECK DANGER	LO	LOCATE OBJECT
WD	WANDER	GO	GRAB OBJECT
ML	MOVE LEFT	LS	LOCATE SUBJECT
MR	MOVE RIGHT	GIO	GIVE OBJECT

STATE	TRANSITIONS					
MF	MF	MR	ML	MB	LO	END
MR	MR	MF	ML	MB	LO	END
ML	ML	MR	MF	MB	LO	END
MB	MB	ML	MR	MF	LO	END
LO	LS	GO	MF	MB	ML	MR
LS	GIO	MF	MR	ML	MB	
LO	MF	MR	ML	MB	GO	

Figure 4.3: The robotic State Machine and the transition table.

Selecting the command corresponding to "grasp" instructs the robot to begin moving to locate the object (with the help of a marker) and evaluate the distance from it. After locating the object, the robot finds the shortest path to the object and begins moving toward it.

When it is near the object, the robot switches to the grasp phase to grasp the object. Similar states have been implemented for the "give" command.

4.2 Assessment of Motivation

4 Healthy controls and 4 ALS patients took part in the study. To asses participants' motivation, a short neuropsychological battery including Mini Mental State Examination (MMSE) [154], Frontal System Behavioural Scale (FrSBe)[155], Neuropsychiatric Inventory, ALS Depression Inventory (ADI)[156] and Beck Depression Inventory (BDI) [157] was administered to all the enrolled participants both ALS (4 participants) and healthy controls (HC) (4 participants). All scores were within normal ranges so cognitive and/or behavioural impairment were excluded. In table 4.1 are reported neuropsychological scores for patients and healthy controls.

4.3 Neuropsychological testing

Before starting the BCI session, current motivation was assessed through an adapted version of the Questionnaire for Current Motivation (QCM, 33). It included 18

	MMSE	NPI	FrsBeS	Beck	ADI
Min	0	0	46	0	0
Max	30	120	174	63	28
Cut-off	24	-	60	14	28
ALS 1	30	4	56	6	13
ALS 2	29.86	0	52	12	21
ALS 3	27.62	2	54	10	15
ALS 4	30	0	48	8	20
HC 1	30	0	48	10	20
HC 2	39	9	59	13	14
HC 3	30	0	58	5	19
HC 4	30	0	46	9	26

Table 4.1: neuropsychological scores for patients (ALS) and healthy controls (HC).

items exploring the four core domains of motivation (*mastery confidence, incompetence fear, interest, and challenge*).

Participants were requested to give a binomial answer (agree/disagree) to each statement (e.g., I look forward to working with the BCI today). The score for each domain was equal to the sum of the affirmative answers on the related statements.

Patients performed the questionnaire through their eye-gaze system while controls used a paper and pencil version. In table 4.2 are shown the scores at the four motivational domains.

The results of the QCM Questionnaire suggest that all participants were highly motivated to use the BCI. Nonsignificant differences were found between the two groups in the four motivational domains.

4.4 Experimental Protocol

This section reports the description of the results achieved by the two groups who took part in the experiment: ALS patients hereinafter defined as ALS and healthy controls, hereinafter defined as HC.

Four ALS (3 males and 1 female), with a median age of 38.5 years and a median education level of 13 years, were recruited. Patients n. 1 and 2 were in a locked-in state, whereas patient 3 and 4 at the time of the experiments preserved residual

	Interest	Mastery confidence	Incompetence fear	Challenge
ALS 1	3/5	3/4	0/5	4/4
ALS 2	4/5	3/4	0/5	4/4
ALS 3	5/5	4/4	0/5	4/4
ALS 4	3/5	4/4	2/5	3/4
HC A	5/5	4/4	1/5	3/4
HC B	4/5	3/4	1/5	2/4
HC C	5/5	1/4	3/5	4/4
HC D	4/5	2/4	1/5	3/4

Table 4.2: Scores at QCM questionnaires for the four motivational domains

arms movements and speech, which was intelligible with repeating.

Four HC (3 females and 1 male), with a median age of 31.5 years and a median education of 15 years were recruited.

In table 4.3 were reported the demographic data and illness parameters of ALS and the demographic data of healthy controls. The healthy control group included four healthy participants not significantly different from ALS patients in age and education.

Participants or their legal guardians signed an informed written consent.

Only one ALS participant had previous knowledge of BCI based system (20 minutes session) while all the other (ALS and HC) were novel to it. All users were trained to limit movement, speaking, blinking and swallowing during the experiment and to report any drawback to the experimenter.

The number of stimuli needed to perform a selection was fixed to 15 visual stimuli to normalize results for all participants.

From preliminary tests, it has been noticed that some of participants obtained good results even with lower number of visual stimuli. ALS patients who took part

Participant	Age	Sex	ALS type	Time since diagnosis (Months)	Artificial	
					Nutrition	Ventilation
ALS 1	40	F	Spinal	180	N	Y
ALS 2	71	M	Spinal	51	N	Y
ALS 3	36	M	Bulbar	48	N	N
ALS 4	26	M	Spinal	12	N	N
HC 1	26	F	N.A.	N.A.	N.A.	
HC 2	28	F	N.A.	N.A.		
HC 3	41	F	N.A.	N.A.		
HC 4	32	M	N.A.			

Table 4.3: The Demographic data and of ALS and healthy controls assessed by the neuro-psychiatrists.

to the experiments were in locked-in states and the commands to be sent to the robot have been defined a priori. A well defined test-bench to test if the action required was correctly accomplished has been setup-up.

Furthermore, using this paradigm, users trained over the system since they had no prior knowledge of BCIs. The experimental procedure is divided in three phases:

- Training Session;
- Online Session;
- Robotic Session.

Each phase was preceded by a detailed explanation of the objectives and details of the experiment. Between each phase a resting session of 5 minutes has been carried out. Between each trial there was a break of approximately 10 seconds to allow user's relaxation.

The *Training Session* was the first phase of the experiment. In this phase, for each subject, the BCI system was calibrated over his or her's neural response to give him full system control.

Accordingly to personal differences, each user ran as many trials as needed to obtain a prediction of 100% of correctness from the classifier.

In average, for all users, 3 trials were needed for the training phase. In this phase no feedback was given to the user.

Each *Online Session* was formed by trials composed by 2 commands per trial: *grasp* and *give*.

Each command was enlighten 15 times. In this phase the user was asked to concentrate on the expected command and, in case of error, the user had to repeat the trial until the correct command was given.

During Online Session, users received visual feedback at the end of each selection, green for a correct command, red for an incorrect one.

In the final phase of the experiment, the *Robotic Session*, the user accomplished 5 trials composed of the same two commands, each displayed 15 times.

In this phase the robot was used as feedback and the user's high level command was translated into a complex behaviour of the robot.

To access the Robot Session, users must achieve an average level of B_f grater than 60% during online session. The overall description of the experiment timeline is shown in figure 4.4.

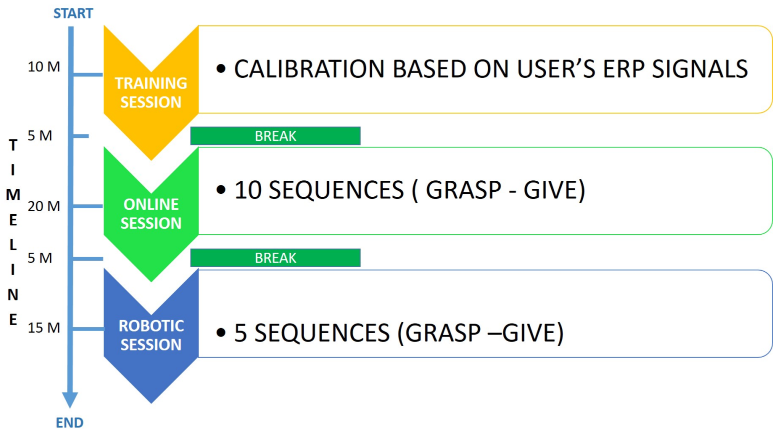


Figure 4.4: the timeline of the experiment.

4.4.1 Environmental setup

The EEG device was connected through USB to the Data PC, which was dedicated to the analysis of signals and management of the user application. The Data PC was connected to two monitors, one is used by experimenters to control and supervise the system, and the other one by the user to run the provided task.

The Data PC was also connected to the BCI server, used to acquire the signal from the EEG device and to process it. The resulting commands were sent, over the network, to a listening NAO server, connected to the NAO robot.

4.4.2 The Results of the experiment

In the **Online Session** the HC achieved a 100% precision in selection task. The average level of Attention was of 74.59% (±5.24%).

The average Central Focus level was 99.035% (±0.82%). The average Intention factor was 43.52(± 3.28%). The overall B_f was 78.15%. The complete list of results is shown in table 4.4.

	Tot Sequences	Correct Sequences	Percentage of Success	Attention	Visual Focus	Intention	B_f
HC1	20	20	100.00%	72.5%	98.05%	37.81%	**75.78%**
HC2	20	20	100.00%	68.20%	99.63%	38.83%	**74.90%**
HC3	20	20	100.00%	78.05%	99.79%	50.53%	**81.2463%**
HC4	20	20	100.00%	79.60%	98.67%	46.93%	**80.69%**

Table 4.4: Online Session results for HC.

The results obtained by ALS, during the Online Session, was 77.083%(± 40.47%). The average Attention was of 69.75%(±15.82%), the central focus of 83.034%(±19.52% and the Intention factor has been of 39%(±6.06%). The overall B_f was 79.61%. The complete list of results is shown in table 4.5.

	Tot Sequences	Correct Sequences	Percentage of Success	Attention	Visual Focus	Intention	B_f
ALS1	20	20	100.00%	86.84%	99.76%	50.91%	**84.82%**
ALS2	20	18	91.67%	75.00%	73.63 %	60.07%	**71.05%**
ALS3	20	2	20%	46.67%	60.84%	66.07%	**56.22%**
ALS4	20	20	100.00%	68.25%	99.05%	80.13%	**82.95%**

Table 4.5: Online Session results for ALS.

Only ALS subject 3 has not been admitted to the Robot Session because he's average biofeedback factor was of 56.22%, under the threshold level set to 60%. Even the percentage of success in selecting the expected command (20%) suggested poor results.

The overall results are shown in figure 4.5, where are shown the results obtained by the HC and ALS (excluding subject 3).

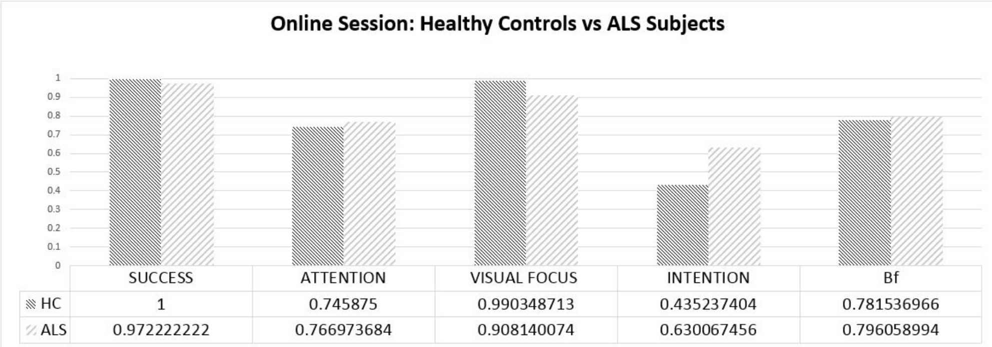


Figure 4.5: HC and ALS results in the Online Session.

Although the sample was not big enough to provide statistical significance, it is possible to derive some interesting consideration from these results.

It is possible to notice how the overall precision for ALS was of 97.22% and it differs only of 2.78% from HC. The Central Focus achieved by the ALS was of 90.81% (with a difference of 8.22% from HC).

Attention and Intention was slightly better for ALS with 76.70% for Attention and 63.57% for Intention factor. The overall B_f value was slighter better for ALS with 79.61%.

During the **Robot Session**, the HC achieved 100% of success over 10 trials, composed by the two command take a glass and give a glass.

The Attention was 69.60%(±8.87%), the Central Focus was 98.49%(±1.97%) and the Intention factor was 42.98%(±12.21%). The list of all values is shown in table 4.6.

	Tot Sequences	Correct Sequences	Percentage of Success	Attention	Visual Focus	Intention	B_f
HC1	10	10	100.00%	75.50%	99.55%	26.91%	**74.34%**
HC2	10	10	100.00%	58.40%	99.54%	60.66%	**75.31%**
HC3	10	10	100.00%	77.80%	99.60%	33.91%	**77.74%**
HC4	10	10	100.00%	77.80%	99.29%	47.73%	**75.94%**

Table 4.6: HC results in robot Session.

The ALS, as shown in table 4.7 achieved 96.97%(± 5.25%) of success over ten trials, each composed by two commands (take glass and give glass).

The Attention was 79.45%(±6.68%), the Central Focus was 96.16% (± 4.82%)

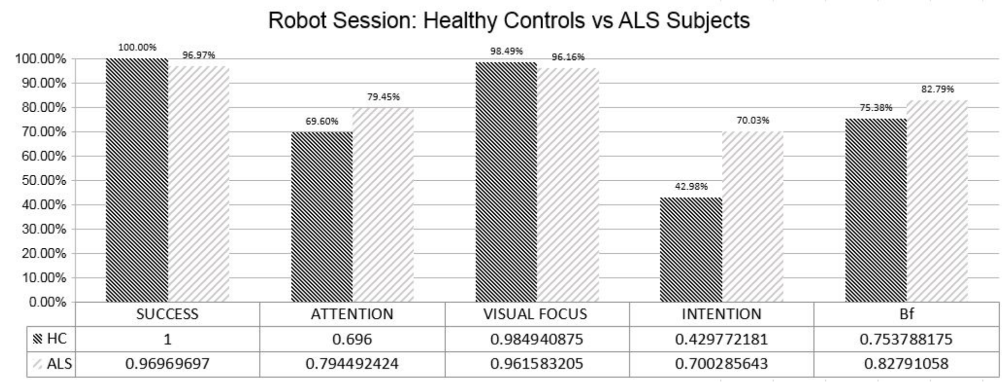


Figure 4.6: HC and ALS results for the robotic session.

and the Intention factor was 70.03%(± 13.49%). The overall B_f value was slighter better for ALS with 84.25%. The results are shown in table 4.7.

	Tot Sequences	Correct Sequences	Percentage of Success	Attention	Visual Focus	Intention	B_f
ALS1	11	10	90.91%	87.27%	98.64%	51.05%	**84.58%**
ALS2	10	10	100.00%	75.20%	90.61%	82.27%	**82.58%**
ALS3	10	10	95.00%	75.88%	99.23%	77.76%	**85.59%**

Table 4.7: Robot Session results for ALS.

From the comparison between ALS participants and HC it is possible to notice a slightly better rate of success for healthy participants with a difference of 3.03% and Central Focus with a difference of 2.34%.

The Attention was significantly grater for ALS participants with a difference of 9.85% over the HC as well as Intention with a difference of 27.05 percentage points.

Analyzing the factor B_f ALS participants obtained an average value of 84.25% while the HC achieved the 75.83%, so the ALS obtained a better result of 8.42 percentage points. The results obtained from the comparison are shown in figure 4.6.

4.4.3 The analysis of Biofeedback factor

Accordingly to findings, the biofeedback factor B_f has been used to modulate the robot response accordingly a *threshold paradigm* as described in section 4.1.1. ALS participant 3 has been the only one excluded from the *robotic session* because of his low biofeedback factor. In particular, his average B_f during the Online Session was under the threshold level of 60%.

In figure 4.8 it is shown his performance during the Online Session, it is possible to notice that only in trials 1,6,13 and 20 the subject was able to select the correct command and, in these cases, B_f was always over the threshold. Although this result is not supported by statistical evidence, it is interesting to notice that a relation between B_f and correct item selection, could be hypotheses.

In general, for wrong selections, the average B_f was 54.80% ($\pm$4.64%) while the average B_f in correct command selection was 78.17%($\pm$7.89%). These results suggest the B_f was a useful indicator of user's mental states and could refine the ERP based selection paradigm.

The B_f values for each of two phases (*online section* and *robotic section*) were averaged and compared within the two subject classes (*ALS*and *HC*) by paired t-test (figure 4.7) due to the limited number of participants. The results of Online Session were not significant increased between ALS (M= 0.7961, SD = 0.0747) and HC (M=0.7815, SD= 0.0327) [p>0.1, t(160) = 1.5950].

By contrast, the results of Robotic Session were significant increased between ALS (M= 0.8425, SD = 0.0153) and HC (M=0.7583, SD= 0.0261) [p<0.0001, t(70) = 22.0216].

In conclusion, over the whole experiment, ALS participants have shown a stronger B_f (81.20%) rather than healthy controls (76.77%).

These results suggest that both ALS and HC achieved good mental control during the online Session but during the Robotic Session the ALS performed statistically better. This conclusion strength the hypothesis that robot is perceived as a positive reinforcement to enhance user's performances, with particular attention to locked-in participants.

4.4.4 The results of the Entropy Module

Accordingly to findings in [101] the variation in entropy could be interpreted as the variation of concentration for the user. In fact entropy could be considered as a measurement of the complexity of the information in the user's brain.

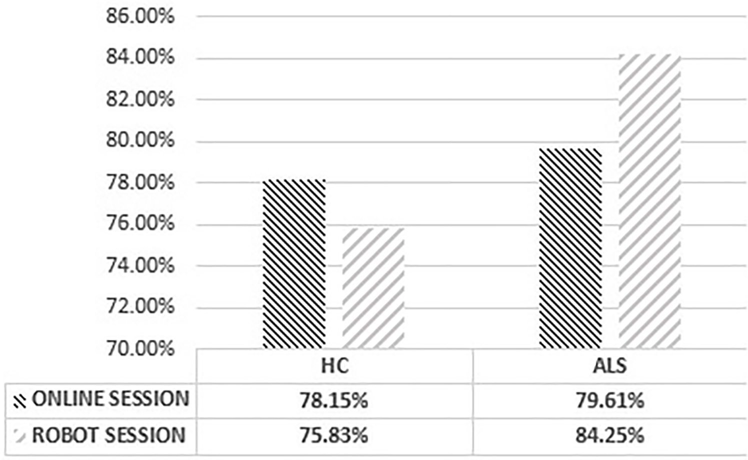


Figure 4.7: HC and ALS Biofeedback factor during online and robotic session.

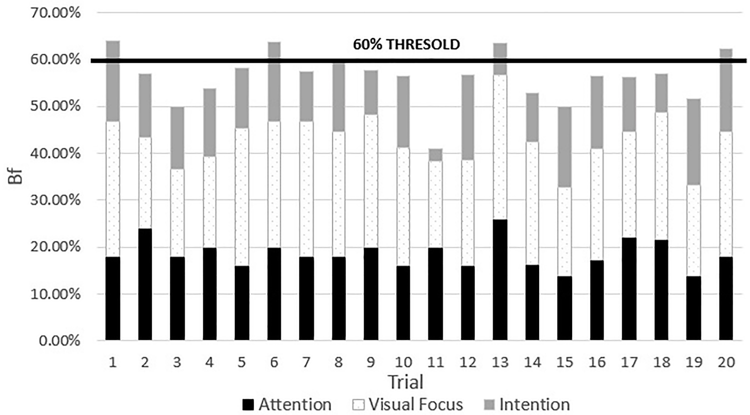


Figure 4.8: The Biofeedback factor trend during the online session for ALS 3 subject excluded from the study. This special case shows that for correct trials (1,6, 13, 20) the B_f overcomes the 60% threshold (dark line).

For this study a post-hoc analysis has been conducted to analyze the values of entropy during the whole experiment. It has been decided to divide the Online session in three parts: *Initial, median and final* and the robotic session in two parts,*Robot Start and Robot End*. ALS have shown a significative reduction of entropy across the experiment while for HC the entropy was not significantly reduced.

The qualitative model, represented in figure 4.9 shows the average variation of entropy for ALS participants and healthy control. Table 4.8 shows entropy values for both group across the experiment.

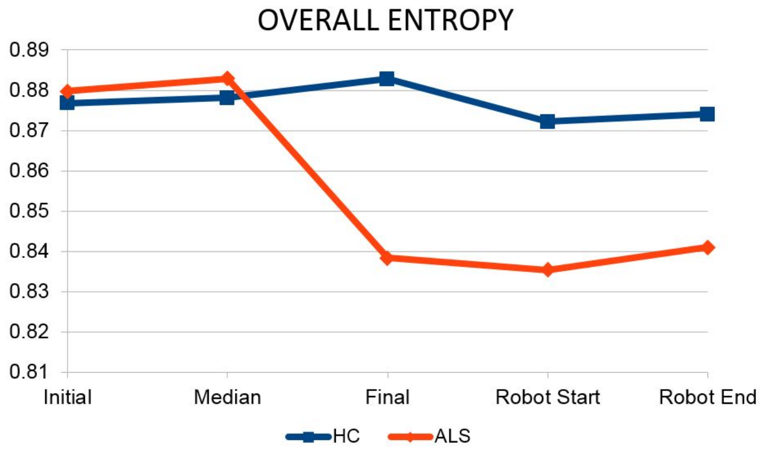


Figure 4.9: The overall entropy during the experiment for HC and ALS participants.

	HC	ALS
Initial	0.8768	0.8798
Median	0.8781	0.8829
Final	0.8828	0.8383
Robot Start	0.8722	0.8354
Robot End	0.8740	0.8410

Table 4.8: Entropy values for ALS and HC during Online Session (Initial, Median, Final) and Robot Session (*Robot Start*, *Robot End*).

4.5 Conclusions

This study presents a flexible BCI-Robot system based on ERPs and biological parameters for locked-in participants to interact with a humanoid robot able to act as user's avatar. All participants were able to control the robot through a BCI interface and it's therefore possible to conclude that the system is able to help a human in accomplishing his needs has stated in research question 1.

The architecture was also able to actively measure the Biological factors providing a measurement of the overall user's mental states (research question 2).

Healthy Controls achieved slighter better results in terms of *Percentage of Success* and *Visual Focus* while ALS achieved better performances in *Attention* and *Intention* and a higher biofeedback factor B_f along the experiment.

Accuracy results were not effected by postural position which varied between participants, since ALS 1, 2 and 3 were bed-bound while HC and ALS 4 carried out the experiment in a laboratory setting. The high accuracy rate and the consequential effective control of the robot.

HC and ALS 4 carried out the experiment in a laboratory setting, whereas ALS 1, 2 and 3 performed the experiment at their own home. So it is possible to conclude that participants' performances were not influenced by posture.

ALS participant 3 was a bulbar-onset patient. Besides the fact that he did not show cognitive impairment on tests, the low accuracy he showed might reflect a relatively low capacity to maintain a specific attention/ability to focus on the task; moreover The low values of this patient on Visual and attention modules probably reflect an impaired high-level functioning of the frontal lobe. This might represent a ground for a further high-level cognitive analyses on other bulbar-onset ALS patients without an overt frontal-lobe related cognitive impairment.

The enhancement of B_f values for ALS and HC in the *Robot Session* is considered as an improvement of participants' mental state. So robot is perceived as a positive feedback for both healthy participants and ALS patients in the use of BCI.

Analyzing results of the *Entropy Module* it is possible to notice a reduction of entropy. This result suggests an increasing in the complexity of users' mental signal.

The presented study validated a new paradigm to interact with a humanoid robot as avatar providing a positive answer to the proposed sub-question 1: Is the NAO Robot able to help a human in accomplishing his needs?). This represents a further step in the studies on Brain Computer Interface where it is is preferred to use monitor interfaces [158] or synthesized speech processor [159], rather then robots.

The experimental sessions have shown as ALS participants reach to manage the system provided with good results as well as HC in terms of performances. Furthermore, stronger **biofeedback** factors have been recorded during the use in humanoid robot control sessions rather than in the Online ones.

The encouraging results suggest that the architecture proposed can be easily employed to boost the robot capabilities and consequently the ALS user abilities.

it is possible, to provide an opportunity to patients affected by severe motor impairment to increment their sense of independence thereby reducing consequently the daily reliance on carers givers.

Not to be underestimated the economic factors that this kind of system can determinate in terms of benefits deriving from reducing assistive needs and the psychological help coming from the restored basic forms of independence.

This study

The future direction for this research will enlarge the sample sizes of patients including people with motor disabilities resulting from other causes(*such as brainstem stroke or spinal-muscle atrophy*) and will expand the range of tasks available through the humanoid robot. The clinical and demographic variables will be correlated with BCI performances of biological user's parameters without a prior knowledge.

Chapter 5

A BCI Robot mediated human-human interaction

In this chapter it is presented a study on human- human interaction mediated the by Telenoid robot. The main goal is to use an assistive social robots architecture designed for being used by people with severe paralysis.

The study presented in this chapter declines the main research question, proposed in chapter 1 as follows: *Is the Telenoid Robot able to mediate a human-human interaction based on BCI?* The study is inserted into the field of Assistive robotics is a new field of Human Robotic Interaction that is a new subfield of robotics that bridges HRI, rehabilitation robotics, social robotics, and service robotics [160]. The architecture has been tested by a user affected by amyotrophic lateral sclerosis (ALS) in a locked-in state.

The system allows the patient to communicate with the stakeholders using a Brain Controlled Interface, based on Evoked Response Potentials (ERP), to express needing, feelings or writing phases. Stakeholders visualize messages sent by the patient on a GUI and use a teleoperated humanoid robot as an avatar of them to extend their physical presence to interact with the patient even when they are away from him.

The study presented in this chapter has been published in the International Robotics and Automation Journal and it's available as open access research item

[1].

5.1 Introduction

Social robots can serve a variety of therapeutically relevant functions, including providing feedback [161], accomplishing tasks [150] [162], extending user presence [163] and help in the management of the ALS disease [163].

Telenoid Robot is a small-size Japanese robot developed from the idea of transferring human's "presence". Telenoid was developed to appear and behave as a minimal design of human features. Its name derived from "Tele" as teleoperation and "noides" as twin, to reassemble, even in the name, how the robot is an extension of the user himself [35].

The ALS is a neurodegenerative disease characterized by progressive muscular paralysis reflecting degeneration of motor neurons in the primary motor cortex, corticospinal tracts, brainstem and spinal cord. Its incidence (average 1.89 per 100,000/year) and prevalence (average 5.2 per100,000) are relatively uniform in Western countries. At the moment there are no treatment for this disease and the only management of ALS is palliative, for this reason it is required a supportive and multidisciplinary approach [164].

In ALS, paralysis is progressive and leads to death due to respiratory failure within 2-3 years. During the latter stage of the disease, patients suffer of a locked-in condition, where their brain maintains all the cognitive functions operating, in a paralyzed body [165].

Brain Computer Interface allows people in locked-in states to communicate with the external world and to control the environment using brain waves, rather than muscles [166]. Among the paradigms presented in literature, it has been used the P300 oddball paradigm for my architecture.

The architecture has been designed to open a new channel of communication between a person affected by ALS and his stakeholders. Furthermore, the system enables the stakeholders to be in a different physical location using the robot as an avatar to give to the ASL patient the feeling of proximity and embodiment [25].

[1]http://medcraveonline.com/IRATJ/IRATJ-03-00068.pdf

The architecture has been tested on 1 ALS patient. Although the sample is really narrow, this is quite usual when working with ALS patients because the relative incidence of ALS [167] and ALSs' personal availability.

5.2 The Architecture

In Figure 5.1 is reported the architecture. The system, based on UnipaBCI framework [168] is composed by a GUI which provides feedback to the patient (from now on defined user) who wears a bio-signal amplifier.

The amplifier acquires EEG signal in real time from user's brain converting in messages. A cloud based network system, interlaces user with robot controller, where the stakeholder (from now on defined operator) visualize the messages. The stakeholder controls the Telenoid robot in tele-operation mode, in such operating mode, the Telenoid robot repeats each movement of the operator and speaks reproducing his voice.

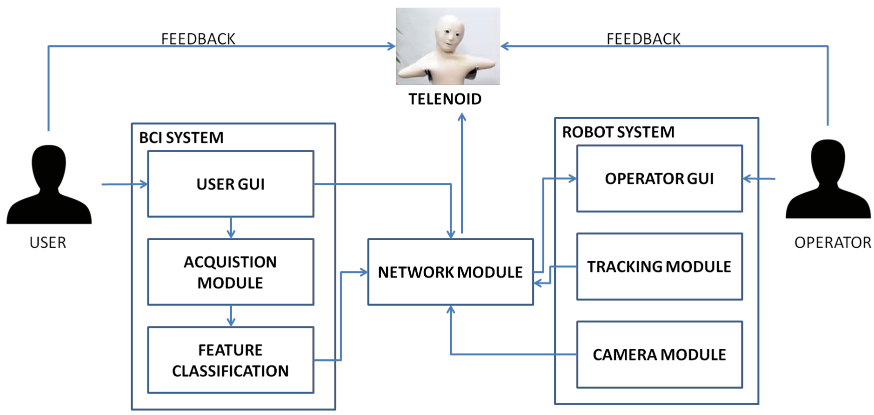


Figure 5.1: The architectural schema of the system

The EEG signals are recorded by the acquisition module, from gel based electrodes placed at four locations of the scalp, according to the 10-20 standard system: P3, Pz, and P4 and Cz. The ground is set at FPz in the forehead and the mass is set on the left earlobe. The signal is digitized at 256 Hz using the g.usb amp, developed by the Guger Technologies.

The signal is then filtered using a Butterworth filter 6th order between 1-40 Hz to reduce artifacts and notched at 50 Hz with a fourth-order notch filter to suppress power line noise.

User GUI, has shown in Figure 5.2 (a), is a 4x4 matrix containing feelings (*Happy, Sad, Content, Embarrassed, Surprised, Angry, Tired, Annoyed*) and phrases (Thank you, Turn the light off, feeling cold, Move me, Turn PC on, I'm hungry, I need you, I love you).

Both feelings and phrases has been chosen by the user during a preliminary screening, selecting them from a list, as the most frequent phrase and feelings he wanted to express.

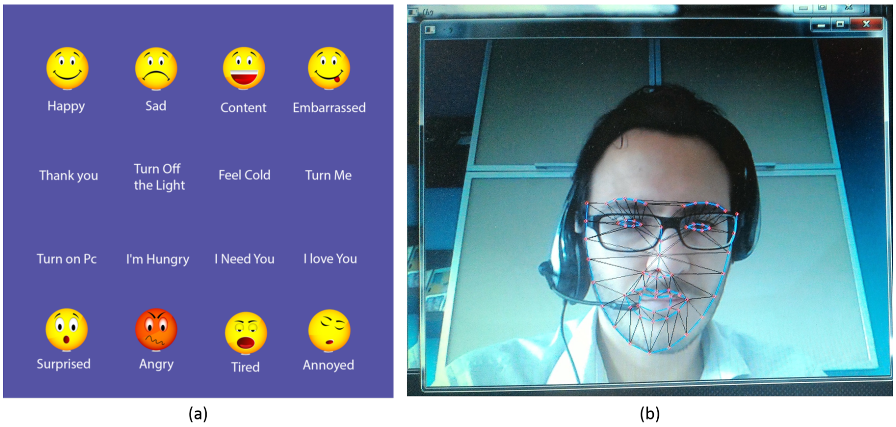


Figure 5.2: (a) User Interface. (b) Operator Interface.

GUI icons are highlighted 20 times (trials) using the oddball paradigm and synchronized with the sampled EEG signals. User is request to mentally count each time the icon of his interest is highlighted. At the end of each trial, the feature classification module is used to calculate user's intention.

To use the system user is required to perform a training phase, where he must mentally count each time known symbols is highlighted to train the classifier over user's P300 response.

The obtained windows are labeled with the corresponding icon and one versus all stepwise linear discriminant analysis [169] is iteratively applied to each window, considering on window as target class and the others as no-target. At the end of this phase, the class with the highest value is selected as target and the corresponding message is sent via the network module to the Operator GUI.

The Network module implements a Control Protocol/Internet Protocol (TCP/IP) and it acts as a middle ware between the BCI system and the Robot System. The Operator Interface, described in figure 5.2 (b) shows the messages sent by the user.

The robot system is natively equipped with a facial tracking system for teleoperation, whose details can be found in [35] and for each operator's head movement, the tracking system calculates robot joints movements to replicate them over the Telenoid. In addition, the camera module shows the scenario captured in real time from the Telenoid camera to the operator.

The described architecture has been tested in a real-world scenario in Palermo, via an ALS male patient of 31 years with no prior knowledge of Brain Computer Interface and robotics.

In figure 5.3 is reported the experimental setting. User lies in this room (Room2) with a BCI monitor in front of him and wears a G.USB Amplifier, connected to a BCI server. User can see the Stakeholder room by a cam connected to a secondary monitor.

User controls a BCI system to send messages to the Telenoid Teleoperation Controller, in Room 1, where Operator sits. The Telenoid Teleoperation Controller is equipped with tracking software and audio-video connection with the Telenoid robot, sets in room 2.

Operator answer to the user speaking on his teleoperation console and Telenoid reacts repeating his head movement and speaking with his voice, giving user the feelings to be with the operator.

5.3 Results and discussion

A total of 6 sessions of approximately 45 minutes for each session has been run with only one participant.

System precision and latency has been evaluated. The latency is expressed as time between the end of a BCI selection and the appearance of the corresponding message over the operator console. In table 5.1 are reported the corresponding results.

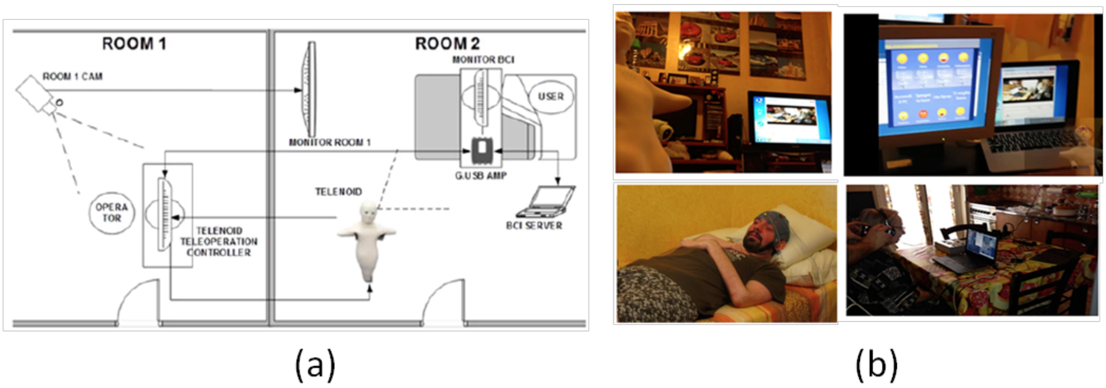


Figure 5.3: (a) a) The experimental setting and (b) some snapshots from the experiment.

ID SESSION	# COMMANDS	CORRECT	PRECISION	LATENCY (s)
1	20	18	0.9000	2.4
2	21	17	0.8095	2.1
3	22	20	0.9091	3.1
4	18	15	0.8333	2.8
5	24	21	0.8750	3.2
6	21	19	0.9048	2.8

Table 5.1: Results in terms of precision and system latency.

The average precision of 90.48 % (± 0.17 %) demonstrates the user's ability to mentally select the desired command.

The system latency is calculated as the composition of the classification time with the network latency. Since the system has been tested over a local Area, the latter could be considered not significant. The observed average latency over the whole sessions is 2.734 seconds (± 0.174).

From the latency it is possible to calculate an average of approximately 22 commands/minutes. The symbolic interface permits to optimize the time for spelling each command, since using a symbolic interface, user does not need to spell each letter of the message he wishes to send but he needs only to select one command.

I administered a simple questionnaire To the ALS patient to measure his attitude toward the system. The questionnaire, reported in table 5.2 was composed only by 3 questions based on 5 points likert scale, where 1 coincided with the minimum and 5 with the maximum. Because of his pathology, the questionnaire was administered using a traditional eye tracking system.

Question	Answer
1) Is the system easy to use	5
2) Was the Telenoid Robot able to express your feelings?	5
3) Are you interested in having this architecture in everyday life?	5

Table 5.2: The Questionnaire

5.4 Conclusions

This chapter describes a novel interaction system for human-humanoid interaction mediated by a humanoid robot as the Telenoid robot was never used for a BCI based communication.

This study is novel in literature, since studies were conducted mainly to investigate BCI bit/rate [170], system precision [171] and user's experience [172]. Moeover, the use of robot in A BCI based communication setting, has been quite neglected since robots were mainly used in controls scenarios [158].

Although the sample is very tight, the results appears to be promising: first, the brain Computer Interface system enabled an ALS patient to communicate using a symbolic interface increasing the brain typing speed. Secondly, the Telenoid robot permitted to communicate, naturally.

Finally, it is possible to conclude that *the Telenoid Robot is able mediated a human-human interaction based on BCI* the Telenoid robot is able to represent human's behavior as suggested by the user's ability to control the BCI system in an accurate way and to express his needs as suggested by the good answer provided by the questionnaire. The main limit of this study is clearly represented by the narrow sample, for this reason it must be repeated with a broader number of users, exploring user's acceptance and satisfaction in a statistic way.

Chapter 6

Exploring humans' acceptance of a neuro-prosthetic Robot architecture

This chapter presents a study on the acceptance of a brain controlled robot architecture. The robot acts as user's neuro-prosthetic extension accomplishing action in his place as human's avatar. The study aims in investigating the research question *Is it possible to model human's responses during human robot interactions, using measurable features?* measuring the acceptance of a human-robot interaction in a real-world scenario.

To support the study, I created an architecture, named A3-K3, for controlling an industrial robot as a neuro-prosthetic extension of a person. The robot used is a KUKA Series 2000 KR 210-2. The robot was fitted with DI which are drawing devices that clamp, hold and manipulate various artistic media like brushes, pencils and pens [1].

User selected an high level task, for instance a shape or movement, using a human machine interface and the translation in robot movement was entirely demanded to the Robot Control Architecture which defined a plan to accomplish user's task.

The architecture was composed by a *Human Machine Interface* on P300, and

[1] http://www.draganilic.org/

by a robotic architecture composed by a *deliberative layer* and a *reactive layer* to translate user's high level command in a stream of movement for robots joints.

The architecture was tested in a real-world scenario, the Ars Electronica Festival 2017 in Linz (Austria) [2], where the A3-K3 architecture has been used for painting. Visitors completed a survey to address 4 self-assessed different dimension related to the interaction between human and robots, for instance the technology knowledge, the personal attitude, the innovativess and the satisfaction.

6.1 Introduction

In the last years BCI technology started to be used to create or modify art-pieces. Users can compose art in real time using brain signals with different paradigms.

Many studies used Brain Computer Interface to create a neuro-prosthetic control systems to stimulates organism to perform arts [173], [174], [175]. In [176] it was tested the possibility to use BCI for creativity expression as part of a software package for entertainment. [177] evaluated the results of a painting application based on brain computer interface on healthy subject and ALS patient. The main limitation of this approaches consists in the use of invasive Brain Computer Interface to achieve devices control. In this chapter I present a novel architecture for using robot as a neuro-prosthetic extension of the user through a not-invasive Brain Computer Interface, evaluating its acceptance from an heterogeneous group of people.

The robot used, Kr-210 series 2000, is produced by KUKA [3], one of the leading manufacturing companies for robotic manipulators in the world. The KUKA Kr-210 could be programmed using two general methods: manual programming systems and automatic programming system with a BASIC-like syntax and simple commands allocated [178].

I based my work on the library proposed by [179] who implemented a cross-platform client to control KUKA Kr-210, using its standard kinematics.

[2] https://www.aec.at/festival/en/
[3] https://www.KUKA.com

6.2 Experimental protocol and scenario

The A3-K3 has been first presented at Ars Electronica Festival (Austria) from the 7th to the 11th September 2017. During the festival the system was used twice per day and a total of ten performances, from now on defined as trials have been held. Each trial lasted 30 minutes and an average of 12 commands per trial were sent to the KUKA robot.

A total of 120 commands has been executed by the robot during the whole exhibition.

To train the HMI a training phase was required. In this phase user was requested to select predetermined symbols from the user interface to train the system over the user's brain response. This procedure required the selection of five symbols for users and lasted approximately 5 minutes. In this phase the Robot Control Architecture provided no feedback.

Once the training phase was completed, user tested the correctness of the training phase selecting at least 3 over 5 correct commands. Finally user was ready to control the robot using the A3-K3 architecture.

The experimental scenario is described in figure 6.1. User was standing in front of the KUKA robot, wearing an EEG cap. A 22" monitor has been set at 1 meter from the user. The RCA Interface has been implemented on a dedicated laptop. Two painting areas has been set one over the floor, defined *Painting area 1* and one over the wall, defined *Painting area 2*.

6.3 The A3-K3 Architecture

In this section the architectural concept underling the A3-K3 system and the high level data flow is presented. Successively, the modules forming the architecture are presented.

The architecture, described in figure 6.2 provide an high level vision of the full system. It is composed by two main module, the Human Machine Interface (*HMI*) and the Robot Control Architecture (*RCA*) modules.

The HMI is used to give to the user an interface to control the robot. Signals were acquired using an EEG amplifier and sent to the bci controller of the HMI .

Figure 6.1: The A3-K3 Architecture.

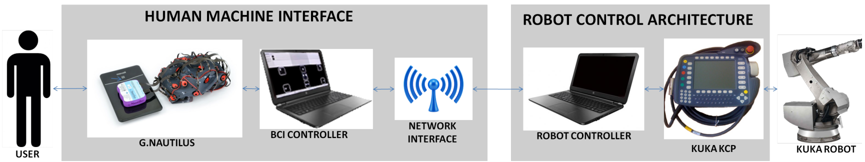


Figure 6.2: The High level description of the system and of its constituents.

The Brain Controller classifies the signals and sends a task over the network system to the Robot Controller Architecture which translated each task in commands for the robot.

The robot used is a Series 2000 KR 210-2. The KR 210-2 robot is typically used for industrial applications. It has a total volume of 55 m^3 and could reach an extension of 2,700 mm. It provides six degrees of freedom and support a payload up to 210 Kg, technical details are reported in figure 6.3.

Robot was equipped with DI Drawing devices that clamp, hold and manipulate various artistic media. For the exhibition pens and pencils have been used, as shown in figure 6.4.

The presented architecture permits to transmit high level tasks to the robot,

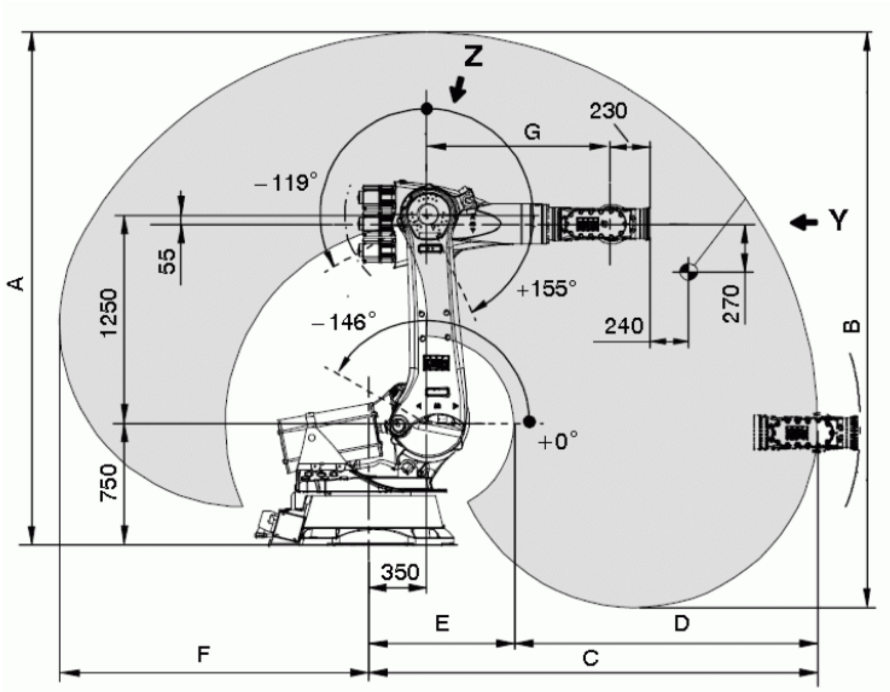


Figure 6.3: The KUKA Kr 210-2 robot, The robot has 6 degree of freedom and roto-translations are defined in forward and inverse kinematics.

which execute them calculating the best plan to accomplish user's intention, reassembling the neuroprosthetic concept of extending own abilities controlling a robot with a Brain Computer Interface.

6.3.1 The Human Machine Interface

The Human Machine Interface is the component of the A3-K3 architecture capable of handling human-machine interactions. The interface consists of hardware and software that allow user inputs to be translated as tasks for the KUKA robot.

In figure 6.5 is shown the Human Machine Interface. The interaction is based on Brain Computer Interface, in particular real time EEG was recorded.

The EEG was recorded using the wireless g. Nautilus (g.tec, Austria) by the *Data Acquisition module.*Electrodes were set in Fz, Cz, P3, Pz, P4, PO7, Oz, PO8 according to the international 10-20 system, reference left ear mastoid, ground Fpz.

Figure 6.4: Some images of the A3-K3 architecture during the Ars Electronica Festival 2017

The data were collected using a sampling rate of 250 Hz and were transferred in real time via Bluetooth to the receiving PC and bandpass-filtered from 1 to 40 Hz by *the pre-processing module*. A BCI wireless cap has been chosen to give freedom of movement during the performance to the user.

To elicit user's response an ERP interface has been used. The interface chosen is based on the Intendix$^{\text{TM}}$ software[4], a commercial software developed by g.tec based on P300 Paradigm. The interface has been customized with icons representing the action that robot will execute. In figure 6.6 is shown the user interface (a) and the task associate to each element of the interface (b).

During each trial, user interface was highlighted using the oddball paradigm. As described in [180] the oddball paradigm elicit an event-related potentials which occurs in the fronto-parietal area of the brain, approximately 300ms after an in-

[4] http://www.gtec.at/Products/Complete-Solutions/intendiX-Specs-Features

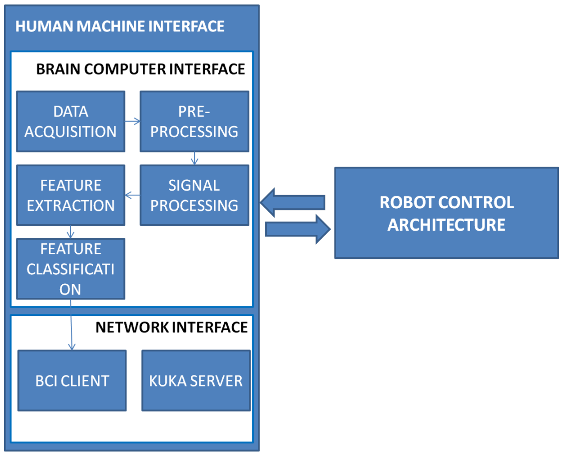
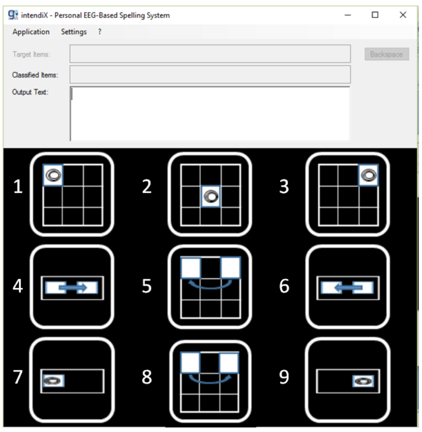


Figure 6.5: The human machine interaction interface. It is composed by a Brain Computer Interface to interact with the user and by a network module to send commands to the robot controller architecture.



(a)

ID	TASK
1	Paint a circle on the left in painting area 2
2	Paint a circle in the center in painting area 2
3	Paint a circle on the right in painting area 1
4	Swipe right on painting area 2
5	Swipe right on painting area 1
6	Swipe left on painting area 1
7	Paint a circle on the left in painting area 1
8	Swipe left on painting area 2
9	Paint a circle on the right in painting area 1

(b)

Figure 6.6: (a) The User Interface. (b) tasks associated to each user interface item.

frequent stimulus.

Accordingly to this founding, frequent and infrequent stimuli are presented to the user, and it is requested to the user to focus on infrequent stimuli, executing a mental process each time an infrequent stimulus occurs. This operation elicit a stronger P300 response [181].

The *Signal Processing Module* mark the real-time EEG with markers which represent an event (infrequent vs frequent stimulus occurred) to synchronize the EEG with the system event. This step is mandatory for the *feature extraction module* which extract a time-locked window to locate the strongest P300 response.

To classify the ERP a one versus all LDA algorithm is applied. The LDA algorithm select one interface item as expected target and considers the other as no-target. In this way the problem to classify the signal is conducted to a two class problem, lowering the complexity. The process is repeated for each item of the user interface and, at the end the one with the higher response is selected as target and sent to the Network Interface.

6.3.2 The Network Interface

The Network interface is used to interconnect the HMI to the RCA. The Network Interface is distributed over the HMI and and the RCA. In fact the BCI module is implemented inside Intendix Software, while the KUKA Server is implemented on the RCA.

Human Machine Interface sends an UDP packet to the BCI Client, which resides on the same machine as the HMI.The UDP interface has been implemented because the command is sent just in one packet. The BCI Client is connected to the KUKA server via a TCP interface. In this way the KUKA robot can be in a different physical location from the HMI.

6.3.3 The Robot Control Architecture

The commands received by the HMI through the network interface is received by the Robot Control Architecture (RCA) as described in figure 6.7 it is based on the model described by [182].

The RCA is a three-level hybrid architecture, which combine a high level component to define plans with a reactive system to ensure system toughness. The highest level is represented by the *World Model* that is used to define a strategy to execute commands.

The *Sequencer* is the middle level and is demanded to the mediation between the plan generated by the world model and the reactive layer.

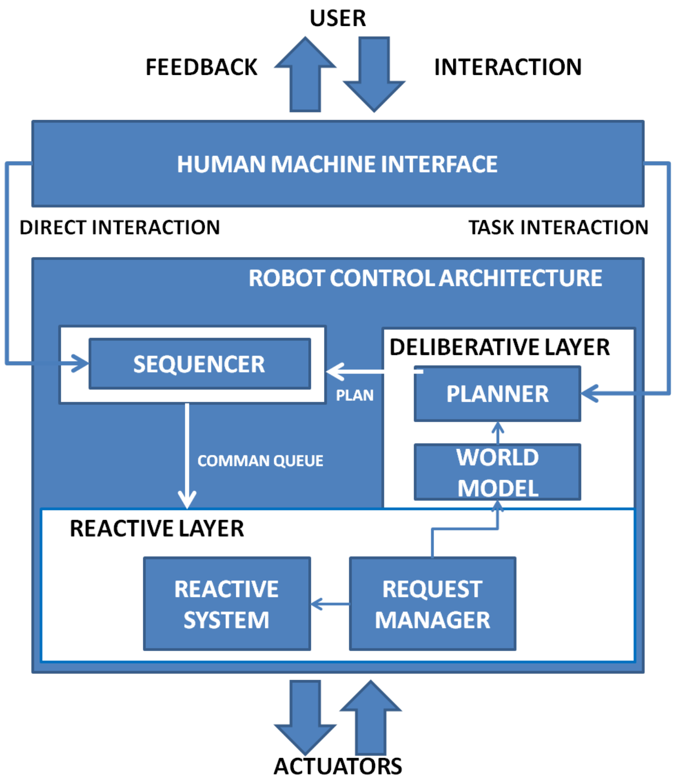


Figure 6.7: The robotic interface receives commands from the HMI calculates robot state by the symbolic layer and translates command in coordinates and position for the joints of the robot. User can also directly control the robot bypassing the symbolic layer.

The *Reactive Layer* is directly connected to actuators in a closed loop fashion. Accordingly to the command received two type of interaction are possible: *direct interaction* and *symbolic interaction*.

Direct interaction is defined as a direct movement command sent to the robot as direction with a fixed length. *Symbolic Interaction* represents an action required to the robot (e.g. *swipe left over the painting area 1*). In figure 6.8 is provided an insight of the module which constitute the Robot Control Architecture.

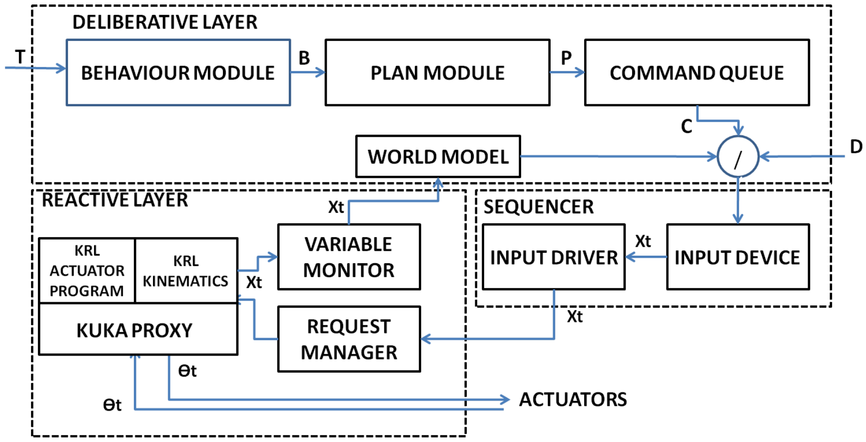


Figure 6.8: The Insight of the subsystems of the Robot Control Interface

The Deliberative Layer

The *Deliberative Layer* represents the highest level of abstraction for the robot. It is composed by several modules to translate an high level command into commands using a planning approach.

The Behaviour module receives a high level command, defined Task T and produce a behaviour B where

$$B = \{p_i \; : \; p_i \, \epsilon \, p_1 \ldots p_k \, with \, k < n\} \tag{6.1}$$

B is composed by a set k of possible n plans P where

$$P = \{c_i \; : \; c_i \, \epsilon \, c_1 \ldots c_k \, with \, k < n\} \tag{6.2}$$

The emergent plan P is obtained from the *planner* and it is decomposed into appropriate "primitive" commands $C \, \epsilon \, c_i$ by the *Command Queue module* to extract the most appropriate c_k from a commands library tailored to the experiment which constitute the basic building block for robotics actions.

To verify if the plan is executable a binary function is defined as If the direct control modality has been set the task is received directly from the HMI and no plan is created. The World model represents the robot position in the environment. As described in figure 6.9 (a), the environment is represented as an mxn grid, and robot occupies a grid cell.

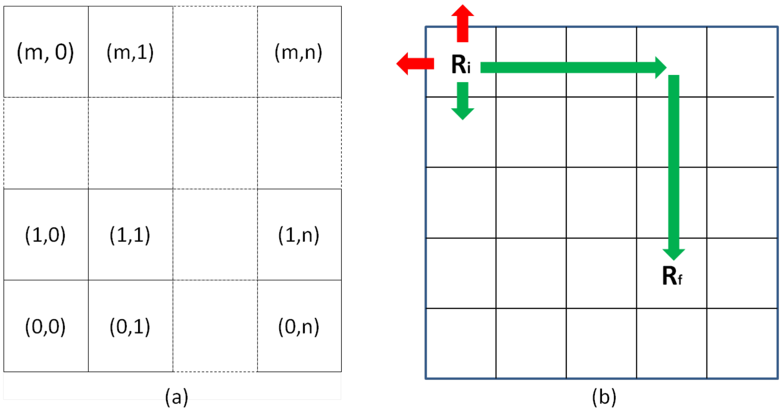


Figure 6.9: (a)The World is represented as a m x n grid. Robot occupies a position on the grid. (b) When a command or a direct control is given to the robot, it evaluates if it is possible to accomplish it or not accordingly to his world coordinates.

As described in figure 6.9 (b), when a new command C or a direct control D is received, it is evaluated the current robot position R_i and the destination position R_f using the following function:

$$F(C,W) = \begin{cases} C & if\, W = 1 \\ 0 & otherwise \end{cases} \tag{6.3}$$

if the action is permitted (in green) , the command is transferred to the *Sequencer*, otherwise no action is performed (in red).

The Sequencer

The sequencer is the module dedicated to produce the low level joints motion for the robot actuators. The *Input Device Module* receives the command queue C producing an output vector which is transferred to the *Input Driver Module* connected to the *Request Manager* module which implements a TCP connection.

The Reactive Layer

The reactive layer has been implemented over the JOpenShowVar Architecture, an open-source cross-platform communication interface, developed by [179]. The input driver sends the command stream to the request manager, connected with the KUKA controller via Transmission Control Protocol/Internet Protocol (TCP/IP). It acts as as a middle ware between the network interface and the KRL.

Since it's not possible to directly set manipulator velocity, using KRC cinematic, speed is expressed as

$$x_t = x_c + x_d \Delta t \tag{6.4}$$

where x_f is the target position, x_c is the current position, x_d is the expected final position and δt represent the time interval between two interactions. The actuators translate x_t a vector containing joint configuration accordingly standard KRL kinematics in the KUKA Proxy.

Actuators states are sent to the variable monitor, which is used to derive current robot position of the robot x_t.

6.4 Results

To evaluate the general attitude toward the A3-K3 architecture and to understand if people perceived such concept acceptable or not, I prepared a questionnaire (Cronbach Alpha 0.8917 Std. Alpha 0.894 G6(smc) 0.9388 Average R 0.5131) which has been submitted to Ars Electronica Festival visitors [5] to assess 4 principal dimension: technology knowledge, attitude, interaction and satisfaction.

The details of people who completed the survey are reported in table 6.1. A total of 681 people (401 Male, 245 Female, 35 N.A), coming from 4 continents (America 19.38%, Africa 5.14%, Asia 12.13%, Europe 57.12%, Oceania 6.17 %) completed the survey. The most representative group age was 30-39 years (33.48%) and 18-29 years (21.59%).

[5]Please refer to this link to access the questionnaire: https://docs.google.com/forms/d/e/1FAIpQLSfFT6wK7Rl p98qnMe2Ec7VksMiAugqV6SBuDJEDUe97Mc2gw/viewform

Age Range	**18+**	**18 -29**	**30-39**	**40-49**	**50-59**	**60-69**	**70+**
Number of people	69	147	228	97	49	48	39
origin	**America**	**Africa**	**Asia**	**Europe**	**Oceania**		
Percentage of People	19.38%	5.14%	12.19%	57.12%	6.17%		
Sex	**Male**	**Female**	**Prefer not to say**				
Number of People	401	245	35				

Table 6.1: People details in terms of age, provenance and sex.

The questionnaire was composed by the following 3-point likert questions:

1. Did you know the KUKA robot?
2. Did you know the Brain Computer Interface?
3. Do you think robots can be used to create art?
4. Do you think the brain computer interface is a useful technology?
5. Was the interaction between the robot and the artist natural?
6. Was the robot an extension of the artist?
7. Was the performance innovative?
8. Did you enjoyed the overall performance?

Questions 1 and 2 refer to technology knowledge, questions 3 and 4 refers to personal attitude toward Brain Computer Interface and Robotic arts, questions 5 and 6 refers to the perceived quality in the interaction, questions 7 and 8 refers to the satisfaction perceived during the performance.

The technology knowledge of BCI and KUKA robots appear to be quite low, in fact only the 20.56% of people had a previous knowledge of the KUKA robot and the 17.33% of people known the Brain Computer Interface where the robot is perceived as a neuroprosthetic extension of the user.

Nevertheless, the personal attitude torward the architecture is well established, since the 59.03% of people considered possible to use the robot for making art and the 58% of people considered Brain Computer Interface a useful technology.

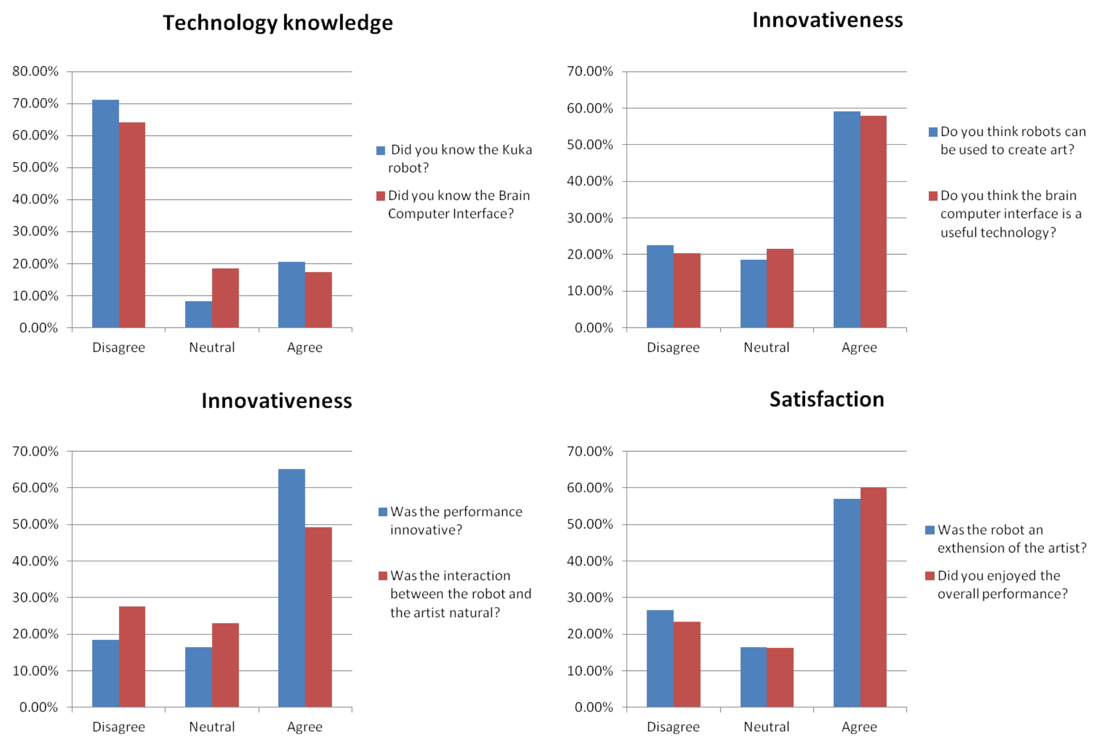


Figure 6.10: Questionnaire results in terms of Technology knowledge, Innovativeness, Personal Attitude and Satisfaction.

The performance has been considered innovative by the 65.20% of interviewed and the 49.34% of people considered the interaction of the robot with the artist, natural. The robot has been considered an extension of the artist himself by the 56.98% of people and the 60% of people was satisfied by the whole performance. In figure 6.10 and table 6.2 it is possible to find the complete list of answer provided to the survey.

6.5 Conclusions

The study presented in this chapter investigates the research question, *Is it possible to model human's responses during human robot interactions, using measurable features?* measuring the acceptance of a human-robot interaction in a real-world scenario over a wide sample.

	Disagree	Neutral	Agree
Did you know the KUKA robot?	485	56	140
Did you know the Brain Computer Interface?	437	126	118
Do you think robots can be used to create art?	153	126	402
Do you think the brain computer interface is a useful technology?	139	147	395
Was the performance innovative?	125	112	444
Was the interaction between the robot and the artist natural?	188	157	336
Were the robot movements fluid and well synchronized?	181	112	388
Did you enjoyed the overall performance?	160	111	410

Table 6.2: The full list of answers to the survey

This study is linked to the emergent question of the scientific community to provide assessment on robots' acceptance [183], [184]. Although acceptance of teleoperated robot has been widely explored, [25], [185], [186], the study presented in this chapter push forward the current state of the art since there are no prior examples of bci-based control of an industrial robot used for art creation.

The presented results provide some interesting insights on people's current knowledge and predisposition torward robotics. In particular, industrial robots like KUKA and Brain Computer Interface are still not well known by the majority of people. Nevertheless people appear to have a positive attitude torward them, accepting them has useful technology in general and in particular to make art.

The performance has been considered innovative by the majority of people and only the 18.36% of people didn't appreciate it.

The most controversial question is about the spontaneity of interaction between user and robot, since it is not considered "natural" by the 49.34% of people.

A possible evolution of the study presented here presented, should be represented by testing the system with the low level command modality, where the users can freely move the KUKA robot on the canvas using BCI. In this was it would be possible to explore if the modification of the command paradigm, will raise the perception of "natural control" in the audience.

Interesting to notice that the 56.98% of people considered the robot as en extension of the artist and they were satisfied by the whole performance.

In conclusions, this chapter presented a novel interaction paradigm based on Brain Computer Interface. The architecture has been designed to be modular with two main systems, the Human Machine Interface and the Robot Control Architecture.

Robot architecture implemented a deliberative layer to create a plan to accomplish a high level command selected by the Human Machine Interface. A low level command modality has been also implemented but not used in presented experiment.

Future study will explore how different type of control will change the system perception. In particular I will investigate if different BCI paradigm and different command modality will assess different acceptance in the users.

Part III

Robot as Teamate

Chapter 7

The analysis of trust in a BCI based Human-Humanoid Interaction

7.1 Introduction

This chapter aims to investigate the role of a robot as a team-mate, investigating the condition under which it is considered *agentic* and *intelligent*. The study declines the main research question into: *Is it possible to derive a model of human's trust from human's brain features during the interaction with a robot?*.

In order to investigate the question, I derived a conceptual model, based on cognitive basis of interpersonal trust, applying it to a human-humanoid interaction.

The conceptual model of trust is derived from the analysis in Heider of the conditions by which participants abide in perceiving and making sense of the social environment [187]. The analysis is simplified in a face-to- face game, where it has been possible to study how the human participants attribute capacity and intention to the robot, *trusting* it as a reliable partner or not.

[188] reports that from a human-human cooperation experiment is possible to demonstrate that an agent's vision of her/his partner's gaze can significantly improve agent's performance in a cooperative task. Similar results were found in competitive game, were a direct correlation between gaze contact with a robot

during a competitive game and the *mentalizing*, intended as the ability to attribute mental states to other agents [189]. Therefore a secondary screen was introduced to take attention away from the robot, to have a measurement of the *saliency* expressed as the inversion of the time participant's had a direct sight contact with the monitor, rather then with the robot; this value asses the *perceptual acquaintance* dimension. [190] and [191] has used the social dynamic of cheating to investigate perceptions of robot agency.

The conceptual model proposed, has been implemented into an architecture, which is an instance of the UnipaBCI framework. To validate the model I realized an experimental scenario, based on a competitive game based on [190] where a robot played against a human occasionally cheating. To assess the trust level I measured some ERPs features, that I considered as the neurological features related to the trust level. In particular I derived such features from P300 and N400. The P300 is related to the context update, while the N400 is related to a violation of expectation. From the obtained results, it appears that these features are spontaneous modulated during the interaction, accordingly to the interaction events. In particular, it is possible to notice a measurable and significative presence of N400 and P300 features during a "cheat condition".

Finally I present the obtained results to provide an answer to the research question of this chapter. The goal is to improve the results obtained in literature, using features extracted from human's real time brain activity to measure the level of trust in a human humanoid interaction.

7.2 The conceptual model

I interpret the *Trust* as the antecedent condition for participants to engage an interdependent relation with other agents.

In this sense trust is not solely making a decision, a choice or the selection of a behaviour like, for instance, cooperating. Trust is modelled as the coordination of independent cognitive modules, which reduce the variability of the changes observed in the social environment by linking them to underlying regularities or invariant dispositions as described in figure 7.1. This definition of trust as well as the model proposed, in figure 7.1 is strictly derived from Heider [192], [193] and

[187] as reffered during further discussion .

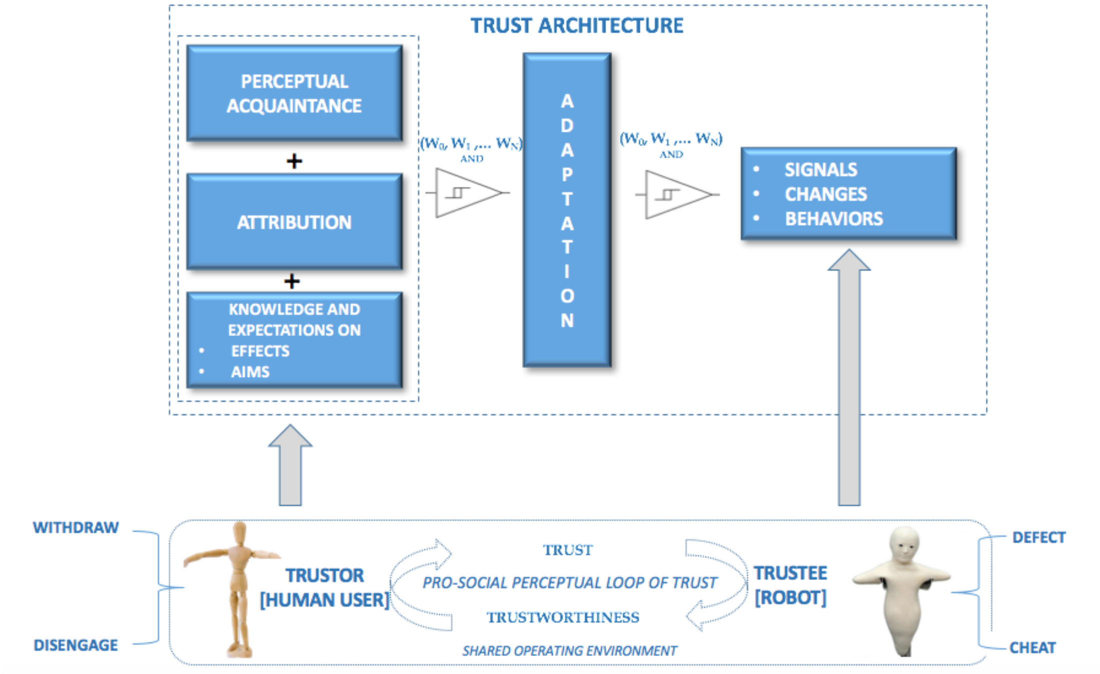


Figure 7.1: The model of trust based on neural features.

According to Heider, agents actions bring about changes in the social environment and introduce variability in the interpersonal relations, which are accounted for by reducing them to the invariance of the dispositions and the intentions of agents that emerge to be reasonably stable across similar conditions [187].

This account requires an implicit perceptual *knowledge* that connects the behaviour of others with *attribution* of another's capacity to execute an observed change, The model represents the knowledge and the attribution as two specialized modules of perception in the social domain.

The interaction takes place indeed in a shared environment where agents are free to move and exchange roles. This does not exclude misunderstanding or delusion in the social domain but constrains them to depend on the expectation and the prediction of behaviour, which turn out to be possible because of the attribution of the effects and changes induced by others' actions to their capacity and intentions.

The proposed model of trust represents the cognitive constituents as three separable modules: the *perceptual acquaintance*, the *attribution* and the *knowledge of the effects* of actions as well as of the aim of the interaction or the task for restricted settings. The information processed by these modules is provided by the overt behaviour of other agents: the actions, or the course of movements selected to carry them out, the changes brought about in the shared environment, intended in the broad sense of encompassing effects on things and other agents, the signals that are unconsciously displayed by agents during the interaction, which have been found to be biologically grounded *honest signals* [194].

The information is collected under the perception of the trustworthiness of agents. The model of trust decomposes the antecedent condition to engage an interdependent relationship in a selected region of the common environment in a pro-social perceptual loop: an agent endows the interaction with trust in another agent who signals her trustworthiness.

The risk and the uncertainty are represented as trustor's withdrawal or disengaging and trustee's defection or cheating. These options amount to the waste or the loss of the resources that are at stake in the interaction and that would have been preserved if agents had remained independent parties in the social environment. An intermediate module represents the adaptation between agents, by which trust can be accounted for as distributed along the continuum within and across people rather than as a single point in all- or-nothing conditions.

The adaptation module serves as weighting function that the three modules apply to the information. If it is given also a threshold function, the model could represent how trust develops during the initial, the enduring and the final stages of the interaction. In this chapter the model is applied to human-humanoid interaction where the human participant is the trustor, the humanoid is the trustee and the region of the shared environment in which the interaction holds is the operating environment.

7.2.1 Biological feedback features

In this paragraph are presented now the ERP features used to measure the level of trust and their contributions. Two features derived from well studied ERPs waves have been taken into account and more precisely, the analysis have been derived from P300 and N400 brainwaves.

The P300 indexes a brain activity when an incoming stimulus causes a change to a particular mental representation and typically it is located in the parietal and central area. In particular, an attention driven comparison process evaluates the representation of a previous event in memory with the representation of the event just occurred.

If no change in the stimulus is detected, the current mental model of the stimulus is kept in memory. For this reason P300 can be defined as a *context update* of a mental representation and his elicited the most by infrequent stimuli or events [69].

Results from Bell et al. [195] proved that the negative deflection of the P300 during a social game outcome not only from the cooperation or cheating but also by their social implications and therefore to the attribution of one player to the opponent.

Starting from these assumptions the hypothesis is that the *cheating* condition, if *attribution* is given to the opponent, should induce a P300 deflection in brain activity therefore P300 is seen as participant's *engagement* in the interaction.

The N400 is negative-going wave that is usually largest over central and parietal electrode sites, with a slightly larger amplitude over the right hemisphere than over the left hemisphere. The N400 is typically seen in response to violations of semantic expectancy and appears to be generated primarily in the left temporal lobe [196].

N400 evidence for interactions of sentence context effects with word frequency, word level associations, and word repetition revealed that higher-level context effects tended to override lower-level ones, contrary to the then-prevailing assumptions of bottom-up priority for and insularity of word level processing [197], [198].

N400 have been found to be a meaning representations of visual real-world events, demonstrating that presentation of the contextually inappropriate information in a film watching scenario, generates such negativity [188].

Starting from these assumptions, a real world scenario has been created, where it is introduced an interaction provided by a human or robot opponent, in which the contextually inappropriate information is represented by the cheating condition which therefore represent a measurement participant's *violation* of expectancy over the *knowledge and expectations on effects*.

The use of features have been introduced in the proposed architecture by the definition of dedicated modules to extract and analyze user's P300 and N400 in real-time, as described in section 7.3.

7.3 The Architecture

The proposed trust architecture is composed by four main functional elements: *Biological Features Extraction*, *User's Action*, *User's Trust* and *Robot Response Controller* and is an instance of the unipaBCI framework [12] as shown in figure 7.2.

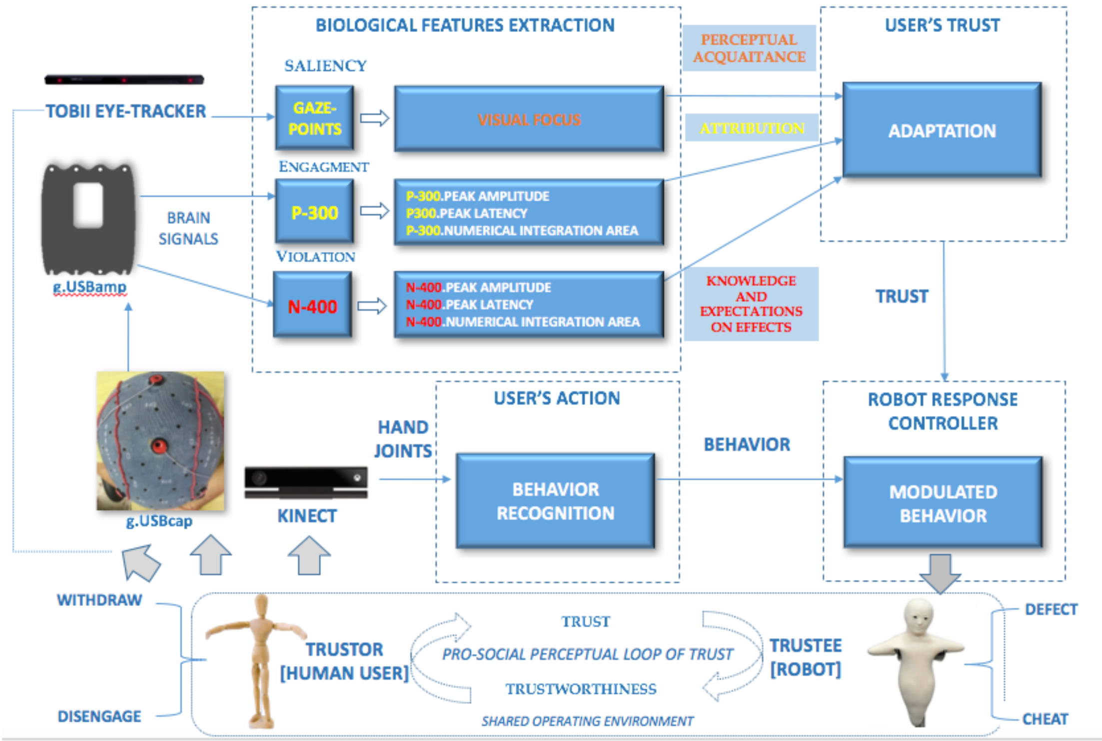


Figure 7.2: The trust architecture.

Biological Features Extraction: In this module, the system concurrently analyses and combines the responses obtained by the g.USBamp brain signals amplifier and the Tobii eye-tracking devices, and provides as output three biological parameters: *perceptual acquaintance* (PA), *attribution*(AT) and *knowledge and ex-*

pectations on effects (KE). In particular, the tobii eye-tracker is responsible of the acquisition of the screen coordinates which the user focuses her attention at.

These gaze points based on their spatial coordinates are used and categorized as visual focus of the user. The g.USBamp bio-signal amplifier provides the raw brain signals produced by the amplifier itself.

From the brain signals obtained by the g.USBamp device, two waves, P-300 and N-400, identified as engagement and violation are extracted. The process of features extraction is delegated to obtain from P- 300 and N-400 three values: the peak amplitude, the peak latency and numerical integration area.

Waveform have been extracted using Erplab, an open source toolkit developed by Lurk et al. [199]. The peak amplitude is defined as the amplitude of a peak over a fixed window and it is expressed in micro-volts μV. The window was set from 250 to 350 ms for P300 and from 350 to 450 ms for N400, to center the window over the expected peak. To avoid noise, waveforms were filtered using as baseline period the 200ms preceding the stimuli, for P300 and the 200 ms following the stimuli for N400 to avoid interference with the P300 waveform. The peak latency represents the latency of the located peak.

From the peak amplitude, I calculated the *numerical integration area*, a value that represents the area amplitude over the selected peak latencies (250-350 and 35-450 ms).

User's Action: The module receives as input the hand joints acquired from the Kinect [1] recognize the behavior exhibited from the user during the game interaction with robot through the behavior recognition module.

User's Trust:this module provides the analysis of the three bio-parameters, representative of the user's inner state. In particular the user's Trust module receives as input three parameters (PA, AT and KE) and produces, using the adaptation module, the current trust level of user during the interaction with the robot. Currently the feedback adaption was deactivated since the main focus was to assess the use of the biological parameters and it is reported just for the sake of completeness.

Robot Response Controller (RRC): this module is responsible of combining together the results obtained by the User's Action and User's Trust blocks.

[1] https://developer.microsoft.com/it-it/windows/Kinect/hardware

The robot, through the RRC, takes the user's behavior and user's trust (currently disabled) as input and provides as output a modulated behaviour.

7.4 Questionnaire

I submitted a questionnaire to the participants after the experimental session to supplement the ERP measures. The questionnaire consists of three constructs derived from the conceptual model and defined to be representative of the cognitive modules of trust. They are represented in figure 7.3 in order to clarify the conceptual relations with the model.

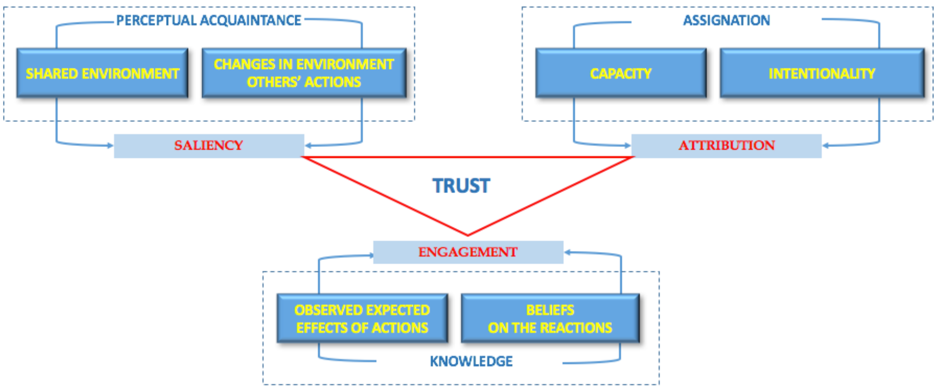


Figure 7.3: Questionnaire constructs and conceptual relations.

The first construct is **saliency** and corresponds to the perceptual acquaintance of the model. The items regard the extent to which the environment is taken to be common to both the participants and the robot and whether the effects of robot actions are attributed to it or to intervening circumstances. The responses collected for this construct measure the embeddedness of agents in the interaction and the conditions at which the robot is perceived as an agent, in particular if its overt behaviour is adequate to give rise to the perception of changes or effects and if it satisfies minimal criteria of consistency without which no such perception could occur.

The second construct is **engagement** and corresponds to the knowledge of the effects on whose grounds participants form expectations robot actions.

The third construct is **attribution**, referring to the human's expectation that the robot is equipped with an intelligent system able to act its choices intentionally rather then to random acts.

7.5 Model validation

The presented model has been validated over humans during a two player competitive game. Starting from the scenario proposed by Short et Al. [190], the trustworthiness of the interaction has been evaluated using human's brain activity. In particular, I evaluated ERPs elicited by external stimuli: the P300 and the N400. The details of this two waves are reported in chapter 2.

7.5.1 Experimental paradigm

The first player is defined as *P1* and represents each person who took part in the experiment. The second player is defined as *P2* and represents the player who bargains over the rules and he has been interpreted by the Telenoid robot opponent and by a human opponent. To test the proposed model two different scenarios, based on the type of opponent that P1 has faced have been set up. Regardless of scenario selection, each game was formed according the following paradigm:

- *Fair Session 1*: Opponent plays following the rules;
- *Cheating Session*: Opponent occasionally bargains over the rules;
- *Fair Session2*: Opponent plays following the rules

The competitive game selected for testing the model was a very simple and well known game: Rock Paper Scissor. This game has been used in a previous work [190] and because it is well-known and simple game.

Rock Paper Scissor rules is a two player game in which each player simultaneously forms one of three shapes with an outstretched hand. These shapes are rock (a closed fist), "paper" (a flat hand), and scissors (a fist with the index and middle fingers extended, forming a V). Scissors is identical to the two-fingered V sign (aka victory or "peace sign). A simultaneous, zero-sum game, it has only two possible outcomes other than a tie: one of the two players wins, and the other player loses.

A player who decides to play rock will beat another player who has chosen scissors, but will lose to one who has played paper a play of paper will lose to a play of scissors. If both players choose the same shape, the game is tied and is usually immediately replayed to break the tie.

Fair Session 1 is formed by ten trials where the opponent plays accordingly to the rules providing audio feedback to the user about the ongoing of the match saying "I won/lose this trial".

Cheating Session 2 is formed by a variable number of trial. In this session, the opponent start to play braking the rules and changing his move every time the cheating condition is verified (cheat to win or cheat to lose). The opponent will also give an audio feedback to the user saying "No, My new move is rock/paper/scissor" and providing information on the ongoing of the match "I won/lose this trial". The session automatically ends when a total of 20 cheating trials have been collected. Fair Session 1 is formed by ten trials where the opponent plays accordingly to the rules providing audio feedback to the user about the ongoing of the match saying "I won/lose this trial".

A total of 2 scenarios have been taken into account:

1. **Human** opponent cheating;

2. **Robot** opponent cheating.

Each scenario differs from the others for the type of opponent: *Human* vs *Robot*. Inside each scenario two cheating conditions have been evaluated: *Cheat to Win* and *Cheat to lose* the zero condition without cheating was defined as *fair*. In figure 7.4 scenarios and conditions are summarized.

7.5.2 The participants

20 participants participated in the experiment with an average age of 22 years ($\pm$ 3 years). All participants were students of the bachelor degree in computer Science 13 male and 7 male. Each participant was randomly assigned to one scenario. All participants were unaware of the cheating session. They were instructed to play

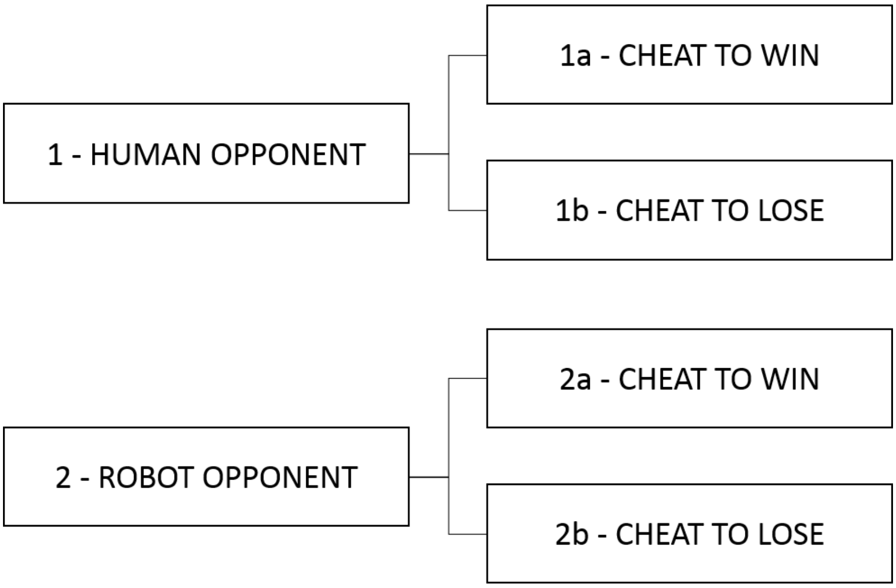


Figure 7.4: The interaction scenarios.

against an opponent at a well known and simple competitive game, namely *Rock Paper and Scissor*, to test their brain response during the game and to test the accuracy of the Kinect system used for tracking their gesture. 10 participants took part to each scenario. Inside each scenario, 5 players played the cheat two win condition and 5 player played the cheat to lose condition.

7.5.3 The experimental setting

The experiment was conducted in a quiet environment, under controlled lighting conditions. P1 sit on a chair at a distance of one meter from P2. P1 wears an EEG cap to track his real time feedback A Kinect was used to evaluate his play move. A feedback screen is placed near P1 but outside his direct sight. On the feedback scren, P1 could see the system gesture recognition and the game score

The experimenter sits in front of P1 and control the experiment with a pc running the control system. In figure 7.5 is reported the experimental setting with P2 robot opponent. For the human opponent the setting is exactly the same and human opponent takes the place of the robot opponent.

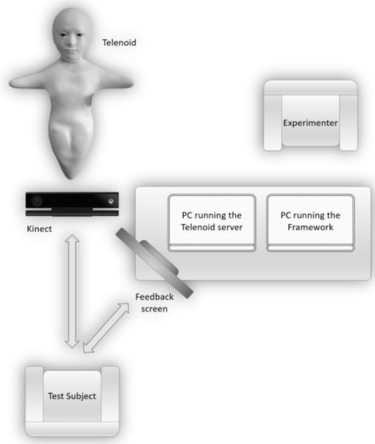


Figure 7.5: The experimental setting.

7.5.4 The electrode montage

The user wears a.gusb amp amplifier with electrodes mounted, accordingly to the 10-20 standard, as reported in figure: 7.6. Electrodes were placed in F8, C3, T3, T5, Pz, C4, T2, T6, P3, P4 O1 and O2. Five electrodes were placed in the frontal area(F1 F3 Fz F4), 5 electrodes were placed in the central-parietal area (C3,C4,P3,Pz and P4), 4 electrodes were placed in the temporal area (T3, T5, T2, T6) and two electrodes were placed in the occipital area (O1,O2).

This montage have been chosen to cover all the principal cognitive areas to asses different cognitive feedbacks in participants brain activity. In particular, for the analysis proposed in the latter part of the chapter only the central-parietal area have been taken into account. The data derived from the remaining areas will be used for further analysis. For each cognitive area taken into account the electrodes of the area were averaged to obtain a response as the mean of the response in the corresponding cognitive area.

7.5.5 Results of scenario 1: Human Opponent cheating during the match

In scenario 1, 10 participants played against a cheating human. 5 participants played along the cheat to win condition (**CW**), the other 5 participant played along the cheat to lose (**CL**) condition. For player, a total of 20 trials were evaluated

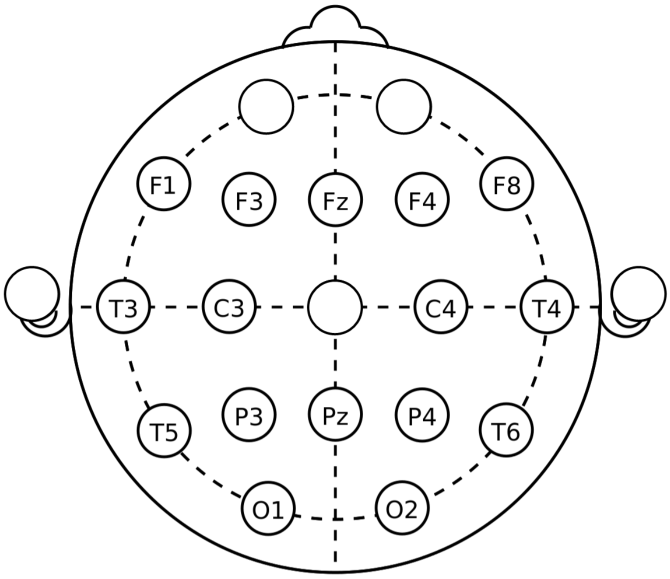


Figure 7.6: The electrode montage.

for the fair condition and 20 trials for the cheating one. A total of 200 Fair trials, 100 CW trial and 100 CL trials have been collected.

Waveform were grand averaged along all trials and participants. A first analysis has been conducted considering a general cheating condition as average of the cheat to win and cheat to lose condition for all participants. In figure 7.7 are presented the difference waves between the cheating and fair condition.

From the figure, It is possible to clearly notice a difference between cheating and fair condition around 300 ms and around 400 ms, suggesting a strong presence of P300 and N400 in cheating condition rather then in the fair condition. From this first visual inspection a careful analysis of ERP features has been conducted.

In figure 7.8 are represented the Cheat to Win and Cheat to Lose condition with their difference waves. It is possible to notice different peaks along these to cheat conditions.

To investigate if that differences could be considered statistically relevant ERPs related features have been extracted and analyzed. In table 7.2 are reported the results in terms of peak amplitude, peak latency and numerical area for P300 and N400.

An independent-samples t-test between Cheat to Win and Fair condition,

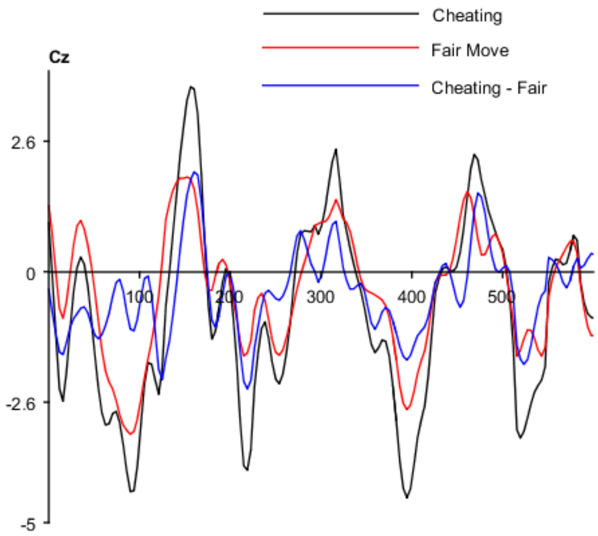


Figure 7.7: The difference waves for cheating and fair conditions playing against a human opponent.

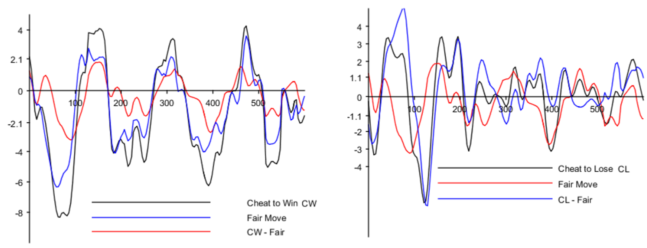


Figure 7.8: The difference waves for Cheat to Win, Cheat to Lose and fair conditions playing against a human opponent.

Cheat to Lose and Fair condition and Cheat to Win and Cheat to Lose conditions for Peak Amplitude and Numerical Integration area for both P300 and N400 features.

The P300 peak amplitude presented a strong statistical significance between CW and Fair condition, $t(298) = 11.2690, p = 0.0001$. There was no statistical difference in CL and Fair condition $t(298) = 1.9119, p = 0.0569$ and a strong statistical significance between CW and CL conditions $t(198) = 8.8177, p = 0.0001$. These results suggest a different biological feedback in Cheat to Win condition, which is not present in Cheat to Lose condition.

	P300 Human Opponent					
	Fair		CW		CL	
	Mean	Variance	Mean	Variance	Mean	Variance
Peak Amplitude	1.436	1.077	3.432	1.991	1.678	0.940
Peak Latency	316.406	12.656	312.500	12.577	320.313	16.016
Numerical Integration Area	0.085	0.035	0.207	0.118	0.046	0.032
	N400 Human Opponent					
	Fair		CW		CL	
	Mean	Variance	Mean	Variance	Mean	Variance
Peak Amplitude	-2.746	-1.126	-6.250	-3.563	-3.130	-1.753
Peak Latency	414.531	11.836	420.625	15.625	422.438	15.938
Numerical Integration Area	0.106	0.043	0.350	0.140	0.073	0.045

Table 7.1: Results features for P300 and N400 in scenario 1.

The P300 Numeric Integration area have been studied to investigate the size of the P300 peak. From the conducted t-test it appeared that there was a statistically significative difference between CW and Fair condition $t(298) = 13.5009, p < 0.0001$ and between CW and CL condition $t(198) = 8.8177, p < 0.0001$. These results strengthen peak amplitude results.

The N400 peak amplitude presents strong statistically difference between Cheat to Win and Fair condition $t(298) = 14.1272, p < 0.0001$ and Cheat to Lose and Fair condition $t(198) = 8.6969, p < 0.0001$. A statistical difference has been found between Cheat to Lose and Fair condition $t(198) = 2.2943, p = 0.0225$.

The N400 numerical integration area scores presents strong statistical significance between Cheat to Win and Fair condition, $t(298) = 22.6361, p < 0.0001$ and between Cheat to Win and Cheat to Lose condition $t(198) = 18.8366, p < 0.0001$. Finally, Cheat to Lose and Fair condition presents a statistical difference $t(198) = 6.1694, p < 0.0001$.

Assuming that the P300 peak amplitude is a representation of the context update of user's mental model and is representative of the effort that participants put in the task, as described in Chapter 2, it is possible to conclude the Cheat to Win condition generated an Evoked Response in participants who played with a human agent.

On the other side, the N400 over the central-parietal area suggest a surprising or unexpected event, as well as a language violation error. Accordingly to the presented results, the Cheat to Win condition generated a stronger N400, compared with Cheat to Lose and Fair condition since the unattended event is probably perceived in participants as surprising or as a violation of their expectancy. For cheat to Lose condition, both P300 and N400 can be identified but their features appears to be lower. This result could be linked to a reduction of the effort put to play with an opponent who plays to lose (P300) and less surprising (N400) rather then an opponent who bargains to win.

7.5.6 Results of scenario 2: Robot Opponent cheating during the match

In scenario 2, 10 participants played against a cheating human. 5 participants played along the Cheat to Win condition (**CW**), the other 5 participant played along the Cheat to Lose (**CL**) condition. For player, a total of 20 trials were evaluated for the fair condition and 20 trials for the cheating one. A total of 200 Fair trials, 100 CW trial and 100 CL trials have been collected.

Waveform were grand averaged along all trials and participants. A first analysis has been conducted considering a general cheating condition as average of the cheat to win and cheat to lose condition for all participants in scenario2. In figure 7.9 are presented the difference waves between the cheating and fair condition.

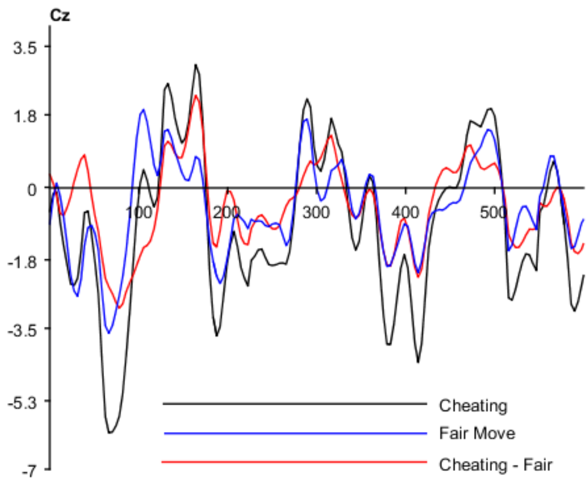


Figure 7.9: The difference waves for cheating and fair conditions playing against a robot opponent.

P300 Robot Opponent

	Fair	CW	CW		CL	
	Mean	Variance	Mean	Variance	Mean	Variance
Peak Amplitude	1.292	0.943	4.012	1.605	0.725	0.580
Peak Latency	316.406	3.164	292.969	8.789	328.125	16.406
Numerical Integration Area	0.062	0.047	0.238	0.117	0.070	0.044

N400 Robot Opponent

	Fair		CW		CL	
	Mean	Variance	Mean	Variance	Mean	Variance
Peak Amplitude	-2.230	-1.160	-8.373	-3.684	-1.566	-0.783
Peak Latency	414.063	8.281	414.063	16.563	418.906	7.578
Numerical Integration Area	0.091	0.063	0.409	0.315	0.027	0.014

Table 7.2: Results features for P300 and N400 in scenario 2.

From the figure 7.9, it is possible to clearly notice a difference between cheating and fair condition around 300 ms and around 400 ms, suggesting a strong presence of P300 and N400 in cheating condition rather then in the fair condition. From this first visual inspection a careful analysis of ERP features has been conducted.

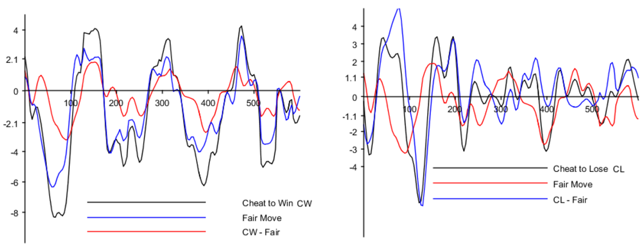


Figure 7.10: The difference waves for Cheat to Win, Cheat to Lose and fair conditions playing against a robot opponent.

In figure 7.10 are represented the Cheat to Win and Cheat to Lose condition with their difference waves. It is possible to notice different peaks along these to cheat conditions. In particular strong P300 and N400 peaks appears to be present in Cheat to Win condition, while no peaks can be located in Cheat to Lose condition.

To investigate if those differences could be considered statistically relevant ERPs related features have been extracted and analyzed. In table 7.2 are reported the results in terms of peak amplitude, peak latency and numerical area for P300 and N400.

An independent-samples t-test between Cheat to win and fair condition, cheat to lose and fair condition and cheat to win and cheat to lose conditions for Peak Amplitude and Numerical Integration area for both P300 and N400 features.

The P300 peak amplitude presented a strong statistical significance between CW and Fair condition, $t(298) = 18.3395, p < 0.0001$. There is a strong statistical difference in CL and Fair condition $t(298) = 8.5539, p < 0.0001$ with a weaker P300 response for CL condition. There is a strong statistical difference between CW and CL conditions $t(198) = 19.2607, p < 0.0001$. These results suggest a different biological feedback in Cheat to Win condition, which is not present in Cheat to Lose condition. Moreover the CL conditions does not elicit a P300 response in participants while playing with a robotic opponent

The N400 peak amplitude presents strong statistically difference between Cheat to Win and Fair condition $t(298) = 21.5696, p < 0.0001$ and Cheat to Lose and Fair condition $t(198) = 5.1640, p < 0.0001$. A statistical difference has been found between Cheat to Lose and Fair condition $t(198) = 2.2943, p = 0.0225$ with a more weak N400 in the Cheat to Lose condition while playing with a robtic opponent.

From these results is possible to conclude that the robotic agent elicited a strong P300 and N400 response during the Cheat to Win condition. No P300 or N400 can be located for the Cheat to Lose condition.

7.5.7 Analysis by type of opponent

In this paragraph is conducted comparison to evaluate if the biological features considered as representative of the interaction with a human (Session 1) and a robot player P2 (Session 2), present similarity.

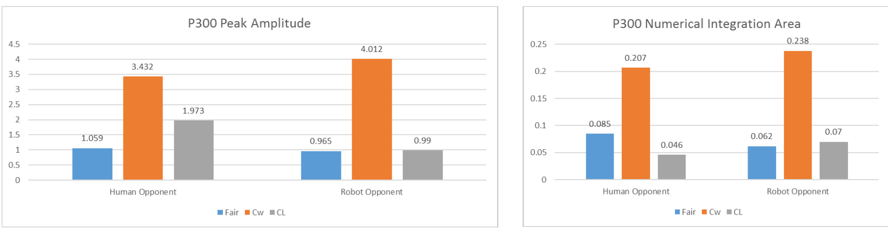


Figure 7.11: The P300 features comparison.

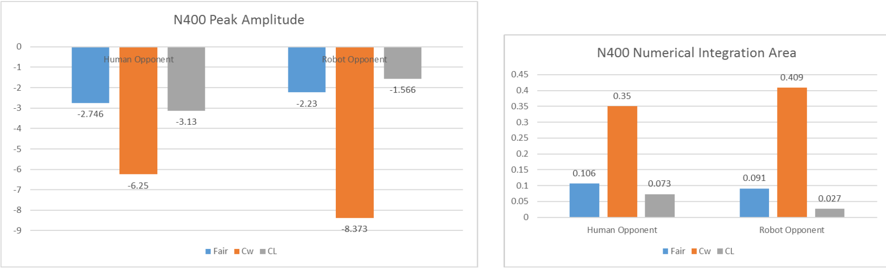


Figure 7.12: The N400 features comparison.

In figure 7.11 and 7.12 are reported a visualization of the P300 and N400 peak amplitude and numerical integration are. Both in human and robot session it is possible to notice a bigger P300 ($t(198) = 2.2680$, $p = 0.0244$). For the cheat to win condition, the P300 peak amplitude was greater in the robot session, rather then in the human one but the numerical integration area was similar ($t(198) = 1.864$, $p=0.0636$). The N400 shown a greater value in the human session, compared with the robotic one ($t(198) = 4.1423$, $p < 0.0001$). The N400 numerical integration area appears to be similar ($t(198) = 2.4548$, $p = 0.0150$).

The fair and cheat to lose conditions have not been taken into account because results from these conditions in Session 1 and Session 2 proved no relevant ERPs components. Nevertheless it is possible to notice a general similar trend for ERPs value both in robot and human session. These results suggest similar evoked response in participants who took part in the study and therefore similar mental attribution to the opponent, considered invariant to the type of opponent but variant to the cheating condition.

7.5.8 The questionnaire contstructs analysis

In this section is presented the results of the questionnaire completed by all participants. Although the considered sample appears to be too small to draw statistical significance conclusion, the reported questionnaire provides an insight of the trend of responses for participants who took part in the experimental session and are treated as qualitative measurement of the interaction with a human opponent and a robotic one along the cheat to win and the cheat to lose condition. All results are presented as average of participants' response along all questionnaire constructs.

To asses questionnaire internal coherence, the Cronbach α was calculated with $\alpha = 0.8675$.

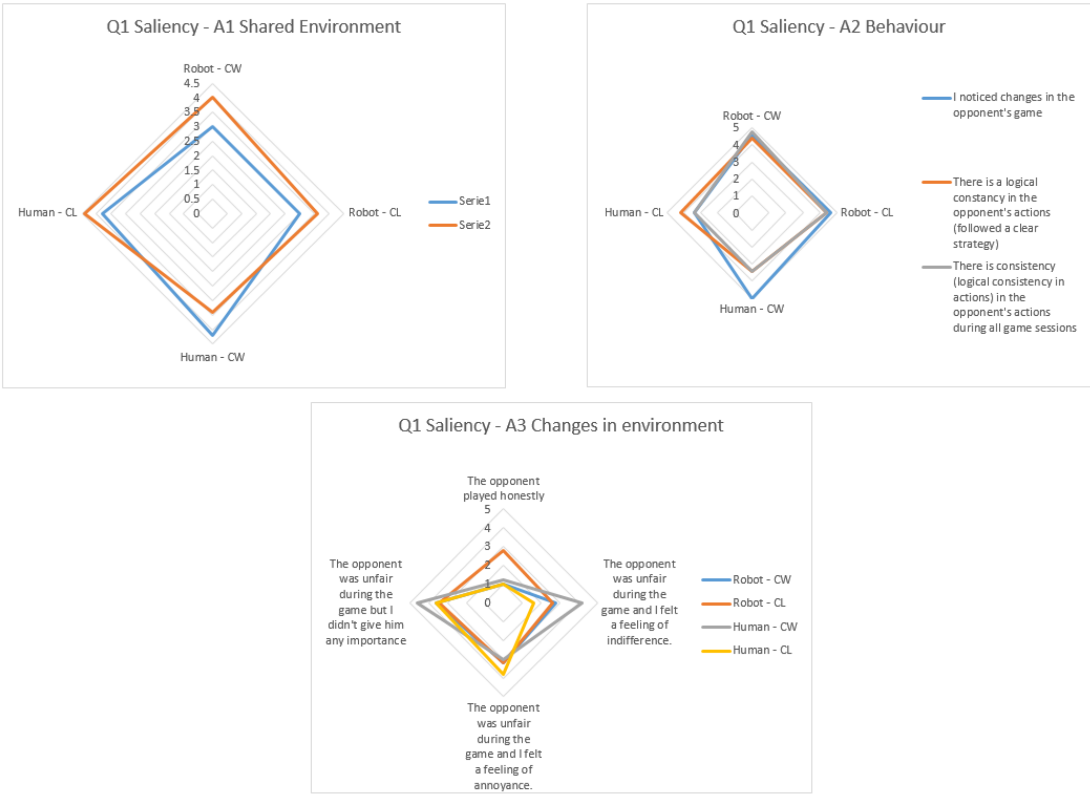


Figure 7.13: Construct Saliency.

In figure 7.13 are presented the results of the saliency construct. It appears that playing with a human opponent gave participants more feeling of sharing an environment (A1). The cheating moves were well recognized in all sessions and conditions (A2). The cheating was found to be particularly annoying when playing against a human an in cheat to win condition for robot opponent session. All participants noticed that the opponents have not played honestly (A3).

In figure 7.14 are presented results related to the engagement construct. All participants focused on the game but during robot opponent sessions, participants appeared to give more attention to the secondary screen (B1). The perceived emotions were mainly surprise and joy with a level of sue fuse for cheat to win condition in robot session and cheat to lose condition in human session(B3).

In figure 7.15 are presented results related to the attribution construct. Answers

Figure 7.14: Construct Engagement

to item D1 shows that the opponent changed his strategy with intelligence both in human and robot session while malfunctions were assigned to the cheat to lose condition. Item D2 proved that under no circumstances the opponent played honestly and the malfunctions are confirmed to be the reason of the cheat to lose condition for robot. On one hand opponents have been being smart in cheat to win condition both in robot and human session while, on the other hand was considered *stupid* during the cheat to lose condition.

The proposed discussion over questionnaire answers along the play sessions and conditions, appears to strength the hypothesis of attribution and engagement during a game with a with a cheating robot, especially in cheat to win condition.

7.5.9 Conclusions

The study presented in this chapter provided an answer to the research question *Is it possible to derive a model of human's trust from human's brain features during*

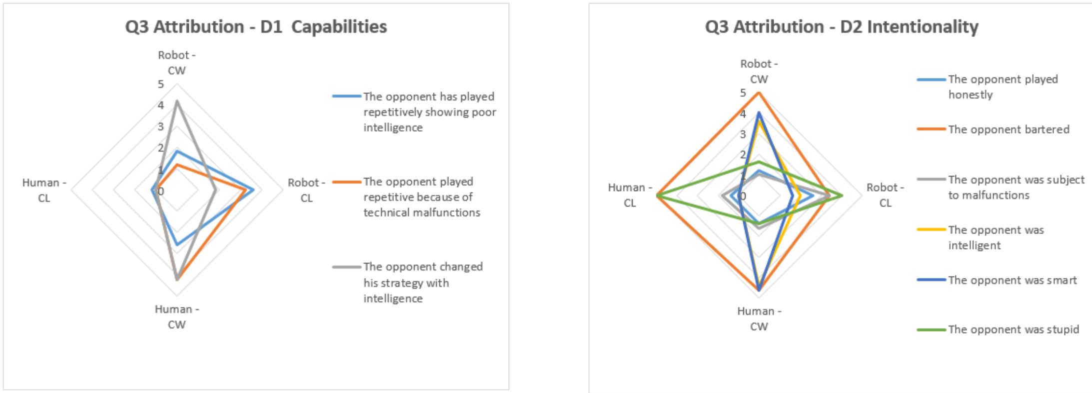


Figure 7.15: Construct attribution

the interaction with a robot? My study starts from the assumption of [190] and [191] who used the social dynamic of cheating to investigate trust and robot agency. The main limitation of the above mentioned researches is represented by the use of speech and gestures to investigate the perception of the robot by the user. The research presented in this chapter provides push forward this limitation introducing an additional measurement of user's trust, by the use of BCI features representing a model based on Heider's researches on trust [187].

In details, during the Human opponent session, it was recorded a well located and high P300 and N400 response, especially during the cheat to win condition, compared with the fair one. These results suggest that the cheat to win condition, being less frequent and not expected from participants it is perceived as an infrequent stimulus able to elicit the Evoked Response Potentials. During the Robot opponent session, similar potentials have been recorded in terms of peak amplitude, latency and numerical area integration.

It is interesting to notice that the responses over the Robotic session was stronger then in Human session. The result obtained for the cheat to win condition with a robotic opponent is in line with Short et Al. [190] conclusion where "the robot is perceived as cunning when it cheat to win" but it is strengthen since the presented chapter provides the biological features to measure the level of mental attribution in robot, and therefore the trust assigned in the interaction, as a combination o biological features. It is possible to conclude that the robot was clearly considered equipped with "intelligence" and able to change its strategy therefore he is considered on the participants' same level and for this reason it is assumed to be perceived as a *Teammate*.

Participants, accordingly to the P300 peak, updated their mental context giving him *attribution* during all the cheat to win condition and the robotic agent was found to generate more ERPs then a human agent. At the same time, playing with the robot, generated stronger N400 ERP in participants. Stated that the N400 is representative of a *violation* of expectancy, the cheat to win condition appeared to violate the participants expectancy (based also on questionnaire construct Q2 - B1). It is therefore possible to conclude that the P300 and N400 represents two component representative of user's attribution and violation, which in the proposed architecture, represent the antecedent condition representative of trust.

On the other hand, the cheat to lose condition did not generate any relevant P300 or N400 in participants during the match with a robotic opponent. This conclusion is perfectly in line with findings of Short [190] who demonstrate how a robot cheating to lose is perceived as not operating correctly. Again the use of the P300 and N400 provided relevant features to define the lower limit to consider a robot as agentic. In fact participants involved in the experiment considered the robot cheating to lose as not intelligent (Questionnaire Q3 - D2). Therefore the robot is no longer perceived on the same level on the participants (a teammate) but it is recognized as a tool with bad functionalities unable to accomplish the expected task, for instance playing in a competitive game with participant.

Finally, using P300 and N400 in a tailored interaction experiment, it was possible to define a lower and upper limit to the trust humans put in the interaction with a robotic agent, to define the amount of "trust" expressed as the antecedent condition for participants to engage an interdependent relation with other agents.

The proposed results suggests that is possible to derive a human's model of trust based on brain features a further study will explore different brain features and social game to investigate if these results could be extended.

Part IV

Robot as Social Mediator

Chapter 8

The role of Honest Signal in a Human-Human Interaction mediated by a Geminoid Robot

8.1 Introduction

In this chapter it is presented a research on user's attitude in a human-human interaction, mediated by a Geminoid Robot considered as a social mediator. The Geminoid (from the Latin "geminus" meaning "twin" and "-oides" that is "similar to") is a teleoperated robot which allows a complete social communication between two people who are physically distant.

The choice of this kind of robot is not arbitrary because the human-like aspect of Geminoid allows various studies on the characteristics of the communication between humans and humanoid robot separating *body* and *mind* as reported by [200].

User's attitude has been measured from honest signalling. As [201] holds, it is likely that the conditions under which human and robotic agents successfully interact and pursue common goals are biologically inspired. The work is based on the theory of honest signalling and sociometrics [123]. The work described in this chapter has been published in the Biological Inspired Cognitive Architecture Journal [202]. The study aims is to explore the main research question (*is it* *possible to model human's responses during human robot interactions, using honest signals?*)

To provide an answer, I created and experimental setting were two users who never met each other must cooperate to solve two priorization tests. One test was solved face to face, while the second one was solved through the Geminoid robot. The order of the two test was randomized. the goal was to monitor the difference in honest signals during a face to face interaction (FFI) and during a Humanoid Human Interaction (HHI).

During both phases I monitored one user's mimicry honest signals (since the other one was then replaced by the Geminoid) through an architecture able to track user's mimicry honest signals. In particular I considered the relative position of one user with respect to his or her interlocutor and mimicry honest signals. The first hypothesis was that during a successful interaction, which was measurable by the evaluation of the scoring of the priorization test, the quantity and type of honest signals should be greater then during an unsuccessful interaction. The second hypothesis was that the Geminoid Robot was the equivalent of a human being for honest signal generation.

The obtained results have shown that the Geminoid robot causes a similar number of honest signals to be generated in humans. The second result obtained is that the priorization task score was higher in HHI rather than in FFI and therefore it is possible to conclude that the Geminoid Robot improve participant's interaction.

The chapter is organized as follows: first I present the proposed architecture, which is an implementation of the UnipaBCI framework. Second, I propose an experimental setting used as test-bed for the architecture proposed in this chapter. Third, I provide a discussion on results.

8.2 The proposed architecture

The creation of this architecture starts from the assumption that the honest signals [203] are a form of separate communication network which support successful face-to-face interactions with clues of reliability and trust among agents. The aim of the architecture is twofold: to serve as an "signalling machine" that employs specifiable

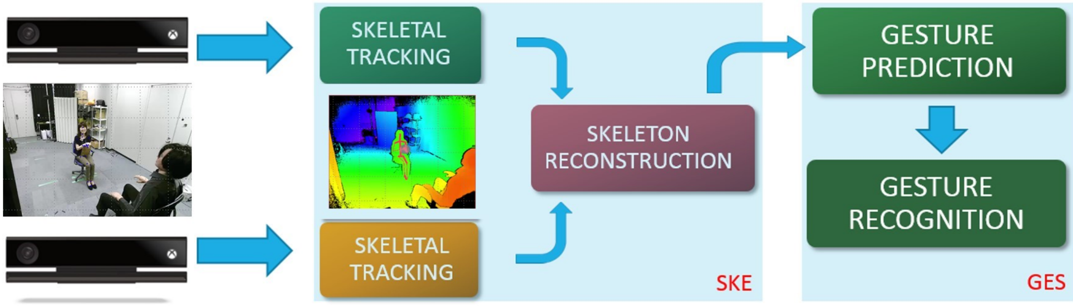


Figure 8.1: The architectural schema.

honest signals as testable parameters in a natural like HRI and to use a Geminoid robot to veicholate such honest signals.

The architecture provides a single shared reference system based on human' posture to have coherent spatial information. The architecture, shown in figure 8.1, consists of two macro modules: *SKE* and *GES*. The *SKE* module use a kinect camera to track the users' skeletal during their interaction with the Geminoid Robot. The module is composed by a skeletal tracking and a skeletal reconstruction block. The first block tracks users' movements over time through the location of their body's joints. The second one, builds the so-called *skeleton* based on the joints obtained from the first one. The *GES* module is composed by a gesture prediction and a gesture recognition block.

In the *Gesture Prediction* the stream of features extracted by SKE modules as skeleton joints are aggregated in a non-overlapping temporal window of 1.5 seconds blocks.The size of the window has been chosen empirically. In the *Gesture Recognition* block a SVM classifier has been trained and used to recognize a set of eight mimicry honest signals: *gesticulating, changing position over the chair, raising folder, lowering folder, touching the head, bending down, writing and holding the folder with one hand*. This actions were selected because they were the most observed during some preliminary, not-structured interactions used only to define which mimicry honest signals should be taken into account.

8.2.1 The Experimental Setup

The experiment took place in a quite room which was ideally divided into two different parts: one for the human-humanoid interaction HHI and the second one for the face to face interaction FFI. A detailed configuration of the experimental environment is shown in figure 8.2. In the room, two Kinect have been positioned for recording the interactions one for the Face to Face Interaction FFI and the other one for the Human Humanoid Interaction HHI. In addition to the Kinect sensors, also two HD cameras have been positioned to record the experiment. Since at this stage the experiment was conducted with people seated, three chairs were positioned in such a way that no occlusions between users or between the Geminoid and the user could occur.

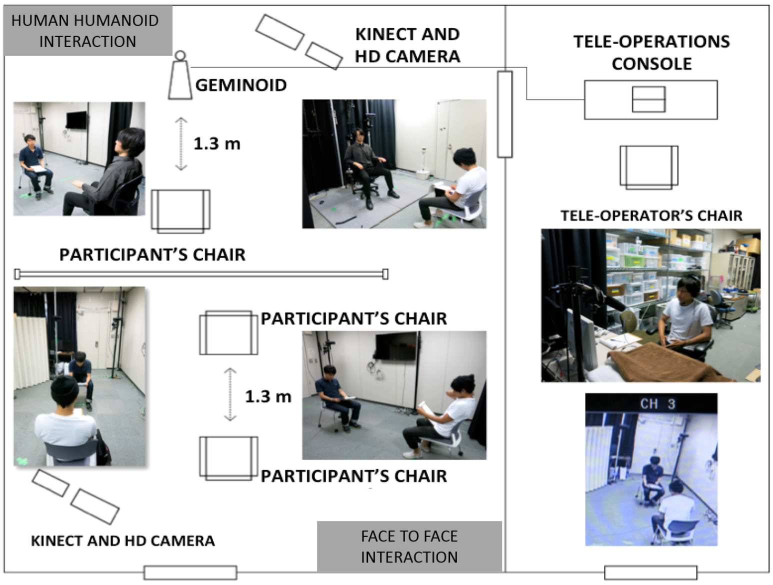


Figure 8.2: The experimental setup.

Twelve paid participants (¥3000 each) have been involved in the experiment and a questionnaire derived from construct proposed by [25] for the robot assessment acceptance was filled before the experiment. The participants' details and their answers are reported in figure 8.1. In figure 8.3 is showed a typical moment of a face to face interaction.

A total of six experimental sessions of *face to face* interaction with the Geminoid

Participants:	
-Number of participants	12
-Male	**7** (58.33%)
-Female	**5** (41.67%)
-Average age	31.33
Previous knowledge or general acquaintance and robotics issue:	
Personal attitude to robotics	
-Real interest	**4** (33.33%)
-Significant knowledge	**2** (16.67%)
-Curiosity	**9** (75%)
-Suspicious	**1** (8.33%)
-Indifference	**1** (8.33%)
Degree of agreement on the Acceptance of Robots in the near Future	
-Accept as useful tools in jobs	**7** (58.33%)
-Accept in all aspects of daily life	**1** (8.33%)
-Accept with suspicion	**1** (8.33%)
-Accept with nuisance	**1** (8.33%)

Table 8.1: The statistics related to participants involved in the main experiment.

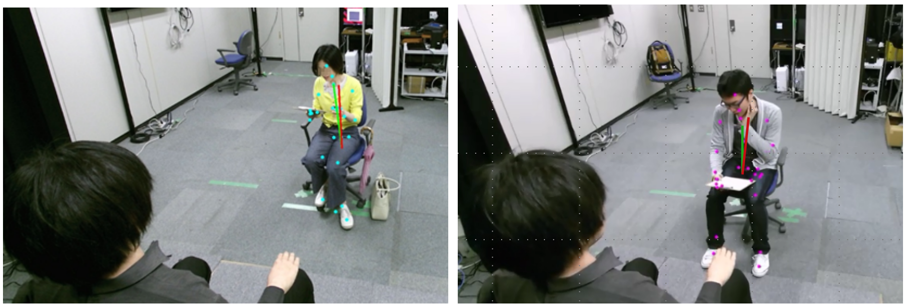


Figure 8.3: Some of the participants involved in the experiment.

robot have been conducted. Before starting the experimental session, each participant has been informed of the session test aims and of its terms and conditions that had been approved by signing. Moreover, have been provided informative documents paper in original language.

The experimental session is started when the people take a seat in the chair placed in front to the Geminoid robot and have an interaction with the Geminoid robot, trying to execute a provided task.

8.2.2 Evaluation of the Interaction task

During the survival prioritization task, the participants were invited to collaborate for solving a problem. In these type of tasks a list of objects is given to each subject and they had to assign to each object a number which decrease with the relevance they attribute to the object. This type of exercise is useful because it forces people to work together overcoming the difficulties in interacting because it fosters trust and friendship among the people who are involved in. The two survival prioritization exercises used in the experiments with the same complexity have been the *Survival on the moon* [1] and *Lost at sea* [2]. They were assigned randomly to HHI and FFI. The participants have assigned, during the survival prioritization task, a value to each object, that decrease with the importance of each of them. The score calculated for each user is compared with a fixed score given by an expert of survival prioritization problems. In order to compute the score, first of all, a subtraction between each value given by the participants and that given by the expert is computed. These differences considered in module are then summed together. The value resulting from this operation is the total score. The evaluation attributed to each score range is proposed in table 8.2. A lower score value indicates a better solution of the provided task.

Score range	Evaluation
0-25	excellent
26-32	good
33-45	average
46-55	fair
56-70	poor
71-112	very poor

Table 8.2: The association between score range and qualitative evaluation.

The time available for solving these tasks were of twenty minutes for each. After each task 5-minutes break has been given to the participants to rest. The results obtained from the answers given by the participants are summarized in table 8.3.

The obtained results shows that in the 85% cases it has been collected a better

[1] http://www.humber.cacentreforteachingandlearning/assets/files/pdfs/ MoonExercise.pdf
[2] http://www.humber.ca/centreforteachingandlearning/assets/files/ pdfs/MoonExercise.pdf

Pair	Score with Geminoid	Evaluation	Score in face-to-face	Evaluation
1	**32**	Good	40	Average
2	**39**	Average	53	Fair
3	**55**	Fair	56	Poor
4	**50**	Fair	52	Fair
5	112	Very poor	**58**	Poor
6	**36**	Average	38	Average

Table 8.3: The results obtained from all the pairs of the experiment.

score in the HHI rather then in the FFI. This result suggests that the use of Geminoid have improved the score during the survival prioritization task enhancing the collaboration between the participants.

8.2.3 The mimicry honest signals Evaluation

Eight mimicry honest signalss were evaluated: (*gesticulating, changing position over the chair, raising folder, lowering folder, touching the head, bending down, writing and holding the folder with one hand*), as reported in figure 8.4.

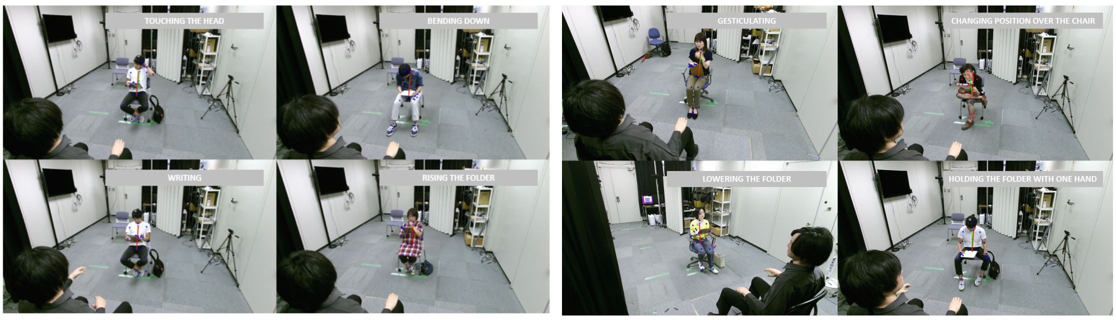


Figure 8.4: The list of all mimicrys honest signals recognized by the system

From the results, I considered only the three mimicry with the most elevated number of occurrences: *rising folder, lowering folder and touching the head* . For each occurrence of the considered mymicrys, I took into account two constitutive measurable elements: the left-right and the forward-backward inclinations of the torso, referred to the coordinate system.

In figure 8.6, are reported the system of reference (A). The joint points chosen for mimicry evaluation (*the yellow points*) (B). The reference system origin, which I set in the *spine_base* joint point of the user body (C). An example of positive

Gesture	Number of Occurrences
Gesticulating	131
Changing position over the chair	49
Rising folder	188
Lowering folder	98
Touching the head	464
Bending down	110
Writing	193
Holding the folder with one hand	13
TOTAL	**1246**

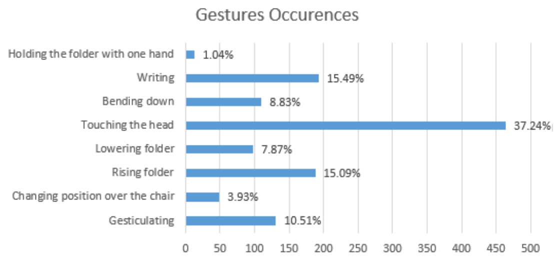


Figure 8.5: The mimicrys recognized by the system.

variation during *forward body movement* (D). An example of positive variation during *left body inclination* (E).

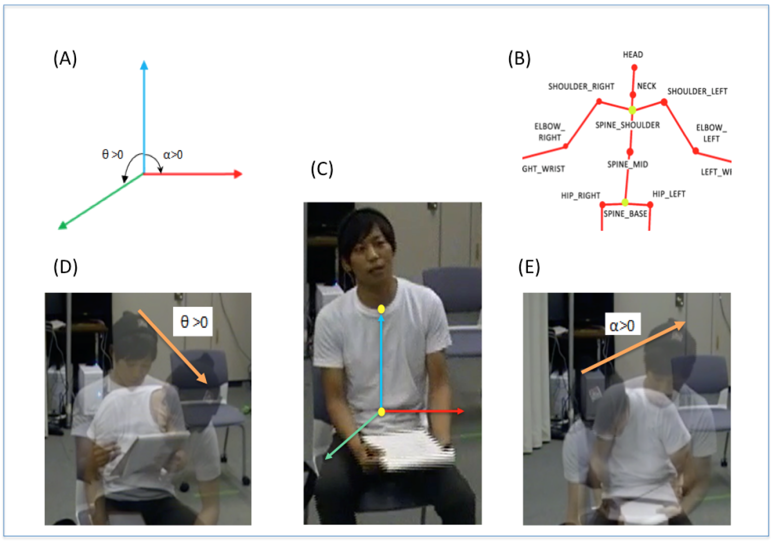


Figure 8.6: Body markers and reference system.

The mimicry honest signals have been evaluated as follows: for each variation of the torso from left to right an α angle is calculated and for the ones from forward to backward a θ angle is measured. The angles were limited from -30 to 30 degrees for the α angle and from -63 to 63 degrees for the θ angle.

Two examples of an angle variation are shown in figure 8.7. All the measurement are expressed as *long sequences*. The term *longest sequences* is referred to the longest movement in the same direction until the user changes the direction of his movement. This was done by counting the number of the forward-backward and left-right moves during mimicrys of each user during the *FFI* and the *HHI*.

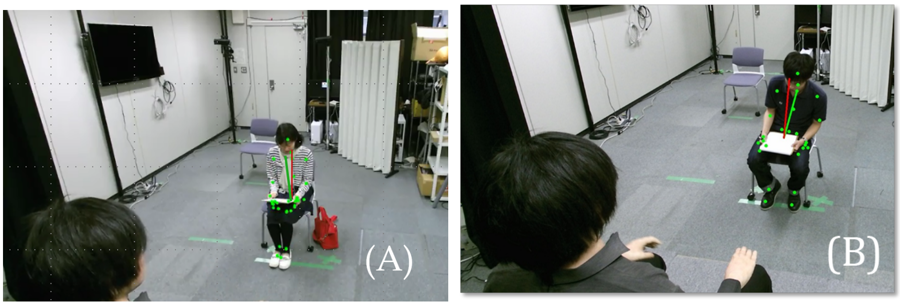


Figure 8.7: Two examples of forward (A) and backward (B) inclination: the reference axis, as *red segment*, and the axis relative to the participant during the movement, as *green segment*, are highlighted.

The average values of the total number of left(LT), right(RT), forward(FW) and backward(BW) variations for each participant involved in both the experimental sessions are summarized in table 8.4.

	FW		BW		DX		SX	
	Mean	Dev	Mean	Dev	Mean	Dev	Mean	Dev
HHI	76.50	41.39	63.00	41.72	50.60	61.06	50.60	61.16
FFI	76.5	47.68	76.7	47.30	56.1	34.16	56.4	34.04

Table 8.4: The average values of the total number of left, right, forward and backward variations and their standard deviation in a range time of 30 minutes.

Accordingly to a Although the sample is small to draw definitive conclusion, these results suggest that the Geminoid robot induced a similar number of mimicry honest signals during the interaction, in fact no statistical differences along all condition (unpaired t-test $\alpha > 0.4873$) could be found in the number of occurrences for forward, backward, left of right occurrences along HHI and FFI and therefore it is possible to consider a Geminoid robot the equivalent of a human interlocutor during a cooperative interaction.

8.3 Conclusions

The study presented in this chapter derive the conditions under which the robot, is able to generate spontaneous signals in a human, during a human-human interaction mediated interaction. The generation of these signal make the robot a social mediator since it is able to provide the base for vehicles human signals in a measurable way.

My study starts from Pentland's honest signal theory [122] that I extend to apply to the research field of human robot interaction. In fact Pentland's study was mainly focused on creating a sociometric instrument to measure human-human interaction [204] while my research goal is to apply honest signals theory to the human human interaction mediated by a humanoid robot.

The study presented in this chapter achieved two results. First assessed that using the Geminoid robot permit to obtain better results during cooperative task between two users who never meet each oter. Second, the Geminoid robot induced in a human a number of mimicrys comparable to those generated during interaction with another human being.

Due to the small sample, no definitive conclusions can be drawn on the assumptions. The study therefore aims to be the starting point for future research on the interconnection between honest signals and humanoid robots.

Chapter 9

Audience interaction with an orchestra mediated by a humanoid robot

9.1 Introduction

In the last decade, intelligent humanoid robots started to be used in common human activities, because of their very "human" aspect. This aspect gives them the ability to convey emotions during the interactions with humans, and this is the main reason that boosted up the development of humanoid robots and their consequent availability for common-life applications. There are a lot of robots that are specifically designed to convey emotions, and their shape is often fine-tuned even on the specific emotional field they are intended for.

For example, emotional humanoid robots were used to sustain natural empathic conversations with elderly people [205], to express emotions and personality [206], to aid people with cognitive disorders [11] and to support disabled people such as "ALS" patients [149]. Furthermore, gestures were suitably used both by humans for the live control of humanoid robots [207], and by humanoid robots to communicate in a "human" way [35].

The hypothesis is that the Human-Humanoid interaction (*HHI*) allows the designing of natural-like settings in which humans and robots coordinate behaviours and emotions [208] and communicate on the basis of shared phenomenal in a common environment [209].

The study starts from the emotional representation proposed by Russel [210], referred to in section 2.8. Starting from these assumptions, I designed a system for conveying audience emotions during a live musical show, using a humanoid robot as an orchestra leader. The goal is to investigate capability of a robot to exhibit affective-expressive movements as an emotional mediator of the feelings of the human audience during the musical performance of an orchestra.

To test this hypothesis, a concert with real musicians was organized in Palermo, in occasion of the Celebration for the foundation of the conservatory Vincenzo Bellini. The concert audience was equipped with a mobile app, which permitted to express their emotion, and to make musicians aware of it. The Nao robot has been used to express the predominant emotion of the overall of the audience to the musicians (sometime referred as orchestra in the following).

The remainder of the chapter is arranged as follows. Section 9.3 includes a detailed overview of the whole system; Section 9.4 shows the first results of an evaluation study conducted during an actual deployment of our proposal; finally, Section 9.5 concludes the chapter and describes future works. The study presented in this chapter has been presented at the International Conference on Complex, Intelligent and Software Intensive Systems [29].

9.2 The Concert Description

Before introducing the architecture, I provide a description of the experimental concert realized in collaboration with the musical conservatory Vincenzo Bellini of Palermo.

The concert was preceded by a preparation phase aimed at generating the music to be played by the orchestra. In particular, some musicians composed a few short pieces, which where processed by a software in order to produce a series of new, original tracks. All the new computer-generated songs, along with the initial ones, were included in the setlist used for the concert. This setlist represent all the pieces that musicians can play. Each track is evaluated in terms of Emotion response accordingly to the emotion recognizer module.

All these tracks were thus sent to the Emotion Recognizer module, which extracted (as explained later) their emotional content in terms of arousal and valence, from here on called *song_arousal* and *song_valence*. The output of this functional block was thus a collection of $\langle$*song_arousal*, *song_valence*$\rangle$ pairs, one per track in the setlist. Accordingly the Circumplex theory, arousal and valence map an emotion.

The concert has been setup with an "act structure". Each act represented an atomic state of the concert. A total of 16 act have been performed, preceded by an introduction and followed by a conclusion. For each act, a total of 4 song extracted from the set-list were selected. Each song corresponded to one of the possible emotions: *happiness, sadness, anger and serenity*. During each act, the audience could express his emotion through a dedicated app. The architecture collected audience's predominant emotion which was used to select the song to be played in the next act by the orchestra. The NAO robot was used to make musicians aware of the audience's predominant emotion.

More in details, during the concert, musicians started by playing the introductory music, while the audience was able to express his emotions using the Emotional Mobile Interface (i.e. *NaoMusic*). All the data sent by this application were continuously stored in a remote Database. When the musical act was close to its conclusion, the *AI Director* queried the Database to fetch the entries gathered during the current act.

Those data were then processed in order to estimate the overall emotional state of the audience, in terms of a $\langle$*song_arousal*, *song_valence*$\rangle$ pair. By using this information, along with the emotional content of the tracks in the next act setlist, the system selected the song whose emotional content best fit the course of the performance, communicating its choice to the Robotic Controller. This module translated the choice into a behaviour for the robot, which informs the orchestra about the next song to be played.

9.3 System Description

In this section, the main building blocks of the architecture are described, as reported in Figure 9.1.

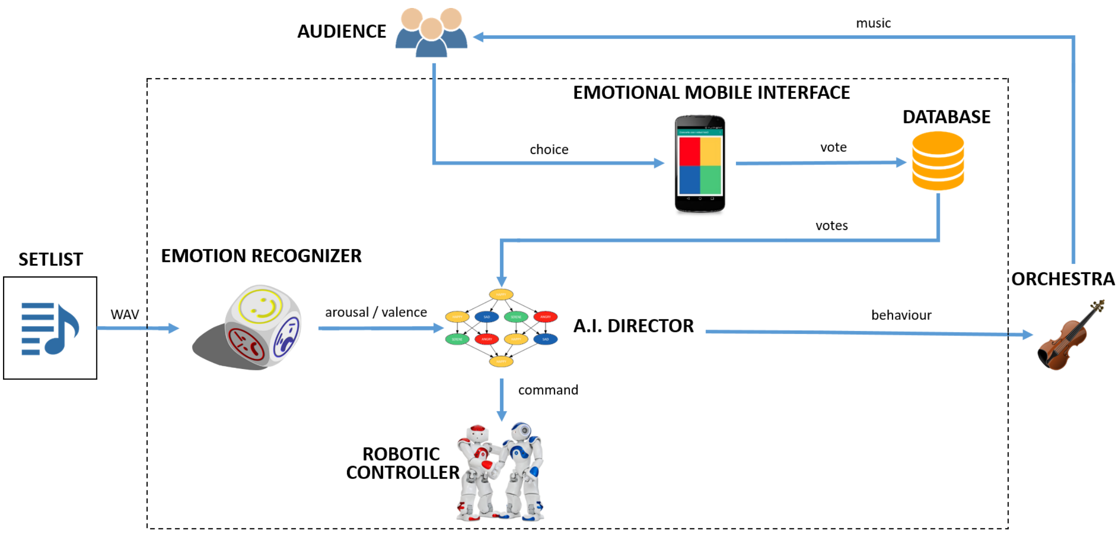


Figure 9.1: System Architecture.

There are three main actors that produce inputs or receive outputs: the *orchestra*, the *audience* and the *setlist*. The first one is made up of an electronic ensemble (i.e. a group of electronic acoustic speakers controlled by a computer) and an instrumental ensemble that produces music with acoustic instruments. The orchestra's role is to receive the system output in terms of gestures from the robot, and to play according to them.

The second actor of this system is the *audience*, i.e. spectators of the concert who are equipped with a mobile application (namely *NaoMusic*). As explained further in the following sections, spectators are able to select their preferred colors according to the emotions felt during the musical performance.

The last actor is the *setlist*, i.e. a collection of audio files that together compose the musical structure of the concert. In particular, these musical tracks are played by the electronic ensemble, acting as the base for the acoustic performance by the instrumental ensemble.

The system architecture is made up of five modules:

- *Emotion Recognizer*: a machine learning module to extract the emotional content from a song;
- *Emotional Mobile Interface*: a mobile app (i.e. the aforementioned *Nao-*

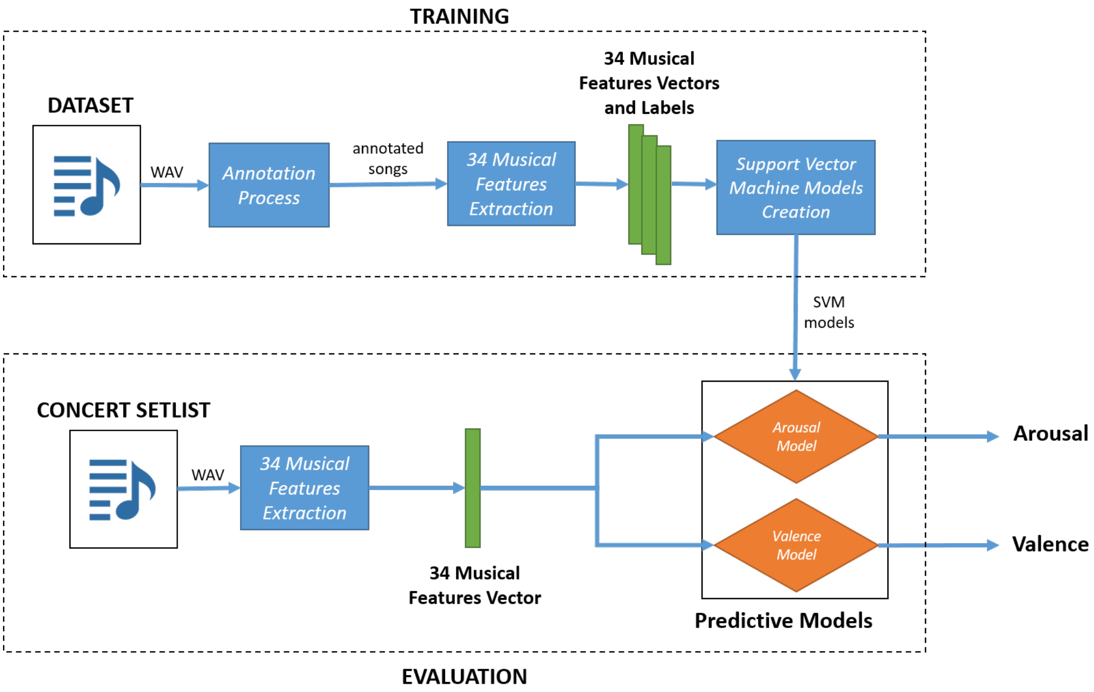


Figure 9.2: Emotion Recognizer module architecture.

Music) that allows users to submit their preferred color, according to the emotions experienced during the concert;

- *Database*: a web-located database that stores all the users' choices;
- *AI Director*: central component that realizes the decision-making process, guiding the evolution of the musical performance;
- *Robotic Controller*: translates the instructions of the *AI Director* module into behaviours for the robot.

Before describing the modules in more details, in the following subsection the use case created to test the interaction between the audience and the orchestra mediated by the NAO robot is presented to give an insight on system functionality.

Name	Size	Description
Zero-Crossing Rate	1	The rate of sign-changes of the signal during the duration of a scheme
Energy	1	The sum of squares of the signal values, normalized by the respective frame length
Energy Entropy	1	The entropy of sub-frames normalized energies. It can be interpreted as a measure of abrupt changes
Spectral Centroid	1	The center of gravity of the spectrum
Spectral Spread	1	The second central moment of the spectrum
Spectral Entropy	1	Entropy of the normalized spectral energies for the set of sub-frames
Spectral Flux	1	The squared difference between the normalized magnitudes of the spectra of two successive frames
Spectral Rolloff	1	The frequency below which 90% of the magnitude distribution of the spectrum is concentrated
MFCCs	13	Mel Frequency Cepstral Coefficients form a cepstral representation where the frequency bands are distributed according to the mel-scale
Chroma Vector	12	12-element representation of the spectral energy where the bins represent the 12 equal-tempered pitch classes of western-type music
Chroma Deviation	12	The standard deviation of the 12 chroma coefficients

Table 9.1: Musical Features used for the SVM-based Emotion Recognition.

9.3.1 Emotion Recognizer

This module implements a method for allowing identification and extraction of $\langle arousal, valence\rangle$ pairs from a musical track. The emotion recognition capabilities have been implemented through Support Vector Machines (SVMs), trained with the data-set described in [211]. It includes approximately 1000 CC-licensed songs that have been listened and subsequently annotated with their values of arousal and valence through crowd-sourcing.

The training process extracts the musical features, shown more in detail in Table 9.1, from each audio track contained in this dataset and uses them, together with the arousal and value annotations, to form the ground truth for a Support Vector Machine (SVM) regression modeling task. More specifically, two models are created: one for arousal and one for valence values. In the evaluation phase, the concert's set-list is analysed and the same musical features of table 9.1 are extracted from each track. These features are then used as input into both the predictive models to obtain the predicted arousal and valence values.

After the training stage, the set-list is processed to extract the aforementioned musical features from each track. These features are then put as input of the two SVMs to obtain the estimated $\langle arousal, valence\rangle$ pairs.

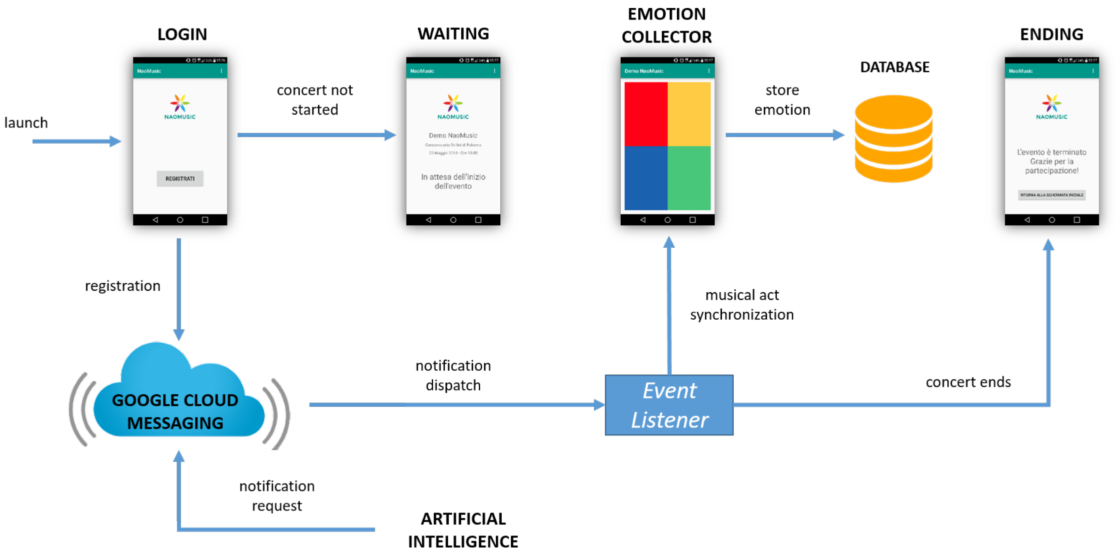


Figure 9.3: Emotional Mobile Interface architecture.

9.3.2 Emotional Mobile Interface

This module plays a fundamental role in the proposed system, as it represents the interaction point between the audience and the entire system. It consists in a simple and user-friendly mobile application, whose interface is shown in Figure 2.8.

Figure 9.3 describes the mobile application architecture. The main goal is to create an interface that allows spectators from the audience for expressing their emotional state while listening to the concert. All the votes recorded by all users at the end of each musical act will be used in order to compute the emotion that more faithfully represents the global emotional state of the audience, and that will be used by the AI Director module to decide which song should be played next.

On start-up, the mobile devices can receive push notifications by the AI Director module through the Google Cloud Messaging (GCM) infrastructure. Once logged in, the application will present users a welcome screen that will persist until a notification will report the beginning of the concert. At this point, the user will be led in the voting screen, where s/he will be able to send the emotion to a remote web server. This interface will synchronize with the musical act progression, thanks to the notification system, to be able to store the votes correctly. This screen resembles Russel's Circumplex Model of Emotions, having four colored tiles

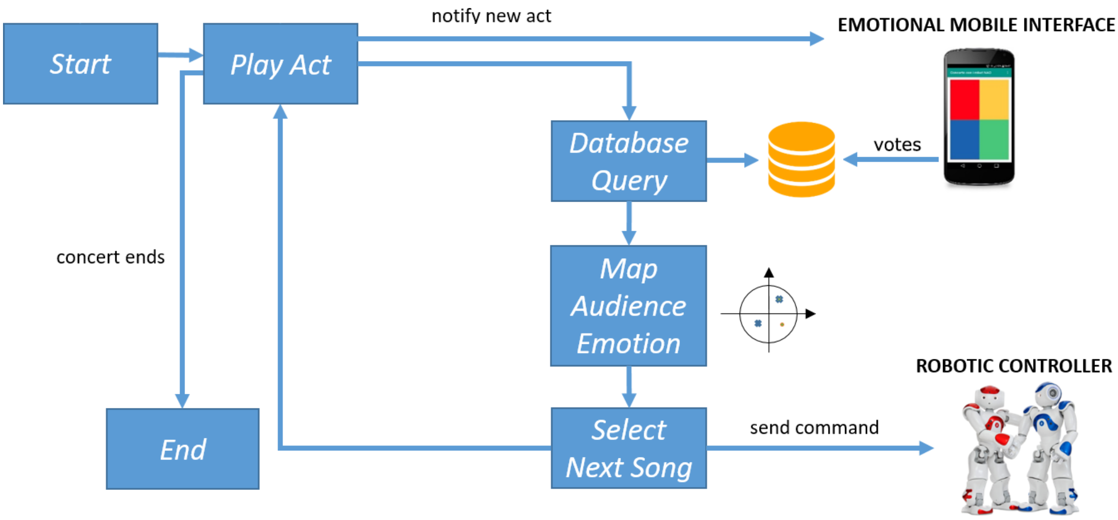


Figure 9.4: AI Director architecture.

that resemble the four emotional categories referred in state of the art (chapter 2).

It is worth noting that nor before neither during the concert users have been advised on the meaning of the colors: they were only asked to interact with the app as they wish, in order to avoid biases and to stimulate a more instinctive and less reasoned response.

9.3.3 Database

The Database module consists in an HTTP server, able to receive GET and POST requests aimed at storing votes to a local MySQL database management system. More precisely, all the votes are firstly sent from the mobile application to a server via GCM, and then stored to the database using REST API. The data are then used from the AI Director module (described in the next subsection), or for further offline analysis.

9.3.4 AI Director

This module is the core component, as it contains the artificial intelligence that deals with the real-time management of the concert, taking decisions on next song to be played. This module accepts as input the $\langle song\_arousal, song\_valence \rangle$ pair

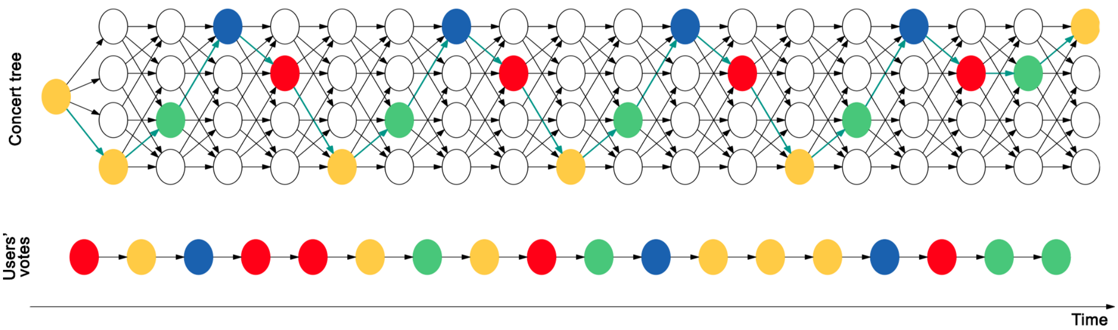


Figure 9.5: The Concert structure. On each act, muiscians can play one of the four possible songs. Each song contains a predominant emotion.

produced by the Emotion Recognizer module.

In order to understand how this module works, it is important to explain the concert structure. It consists in a tree data structure, whose depth represents the number of musical acts, and whose nodes are all the possible songs that can be played in every act (see Figure 9.5). The goal of this module is to find an optimal path from the first to the last node, in order to produce the most engaging effect for the spectators based on their emotional response.

At the beginning of the performance, the robot will perform an introductory welcome animation and subsequently it will turn towards the musicians in order to communicate via a body gesture the beginning of the first musical act.

For every such acts, the AI Director sends a synchronization notification to the Emotional Mobile interface, then it remains in a idle state, giving time to the electronic and orchestral ensembles to perform their executions. Shortly before the end of the act, the system will wake up and query the Database to retrieve the spectators' votes collected during the act.

This information is used to calculate the audience's predominant emotion. Combining this information with the emotional content of the candidate songs eligible to be performed on the next act, a heuristic search algorithm is used to choose the song that most likely would induce an emotional shift in the audience, while avoiding stalls on the same emotions and abrupt changes. For instance, considering the $\langle song\_arousal, song\_valence \rangle$ pair as a point on the circumflex shown in Figure 2.8, its mirrored point (with respect to the main diagonal) represents

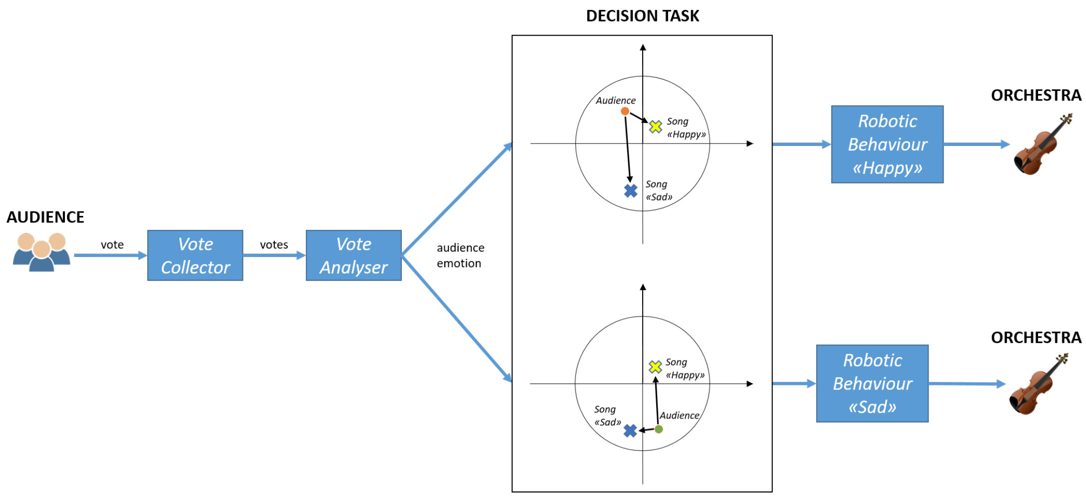


Figure 9.6: The decision process.

the opposite emotion, which can be selected for obtaining the highest emotional variability.

Figure 9.6 shows a more detailed representation of the decision process that occurs at the end of each musical act, with the AI Director module computing the distance between the audience and the songs in the emotional domain.

9.3.5 Robotic Controller

Once the AI Director chooses the next song, the Robotic Controller module will convey this information through the robot. This chain of operations is executed in loop for every musical act.

The purpose of this module, whose architecture is shown in Figure 9.7, is to model the robotic behaviors that will enable a NAO robot to interact through body gestures with people involved in the performance.

The robot plays two main roles: it serves both as an announcer, gesticulating and speaking to the audience, and as a musical director, performing movements to instruct the musicians on the next song they should play. The gestures used for communicating emotions are shown in Figure 9.8, and have been self assessed to be as simple as possible and non-ambiguous.

These four movements are the robotic embodiment of the outputs from the AI

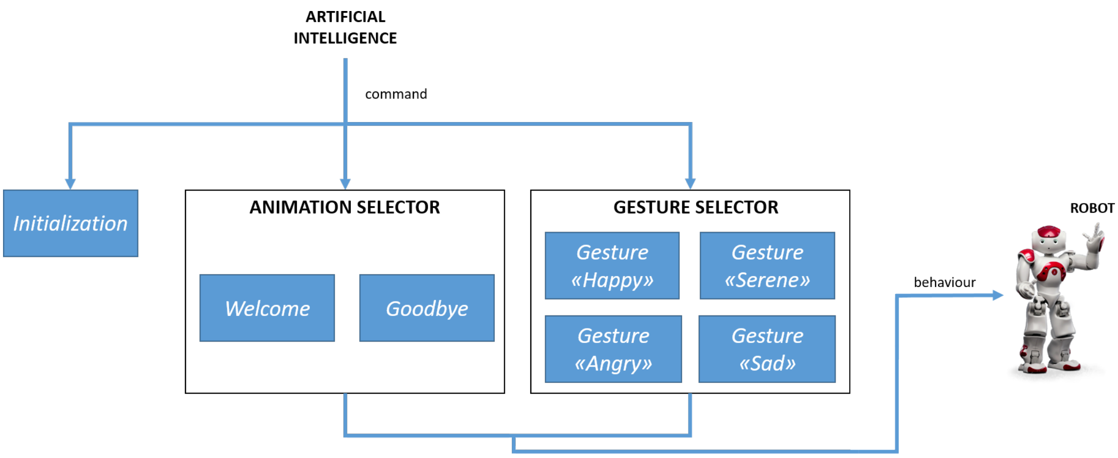


Figure 9.7: Robotic Controller architecture.

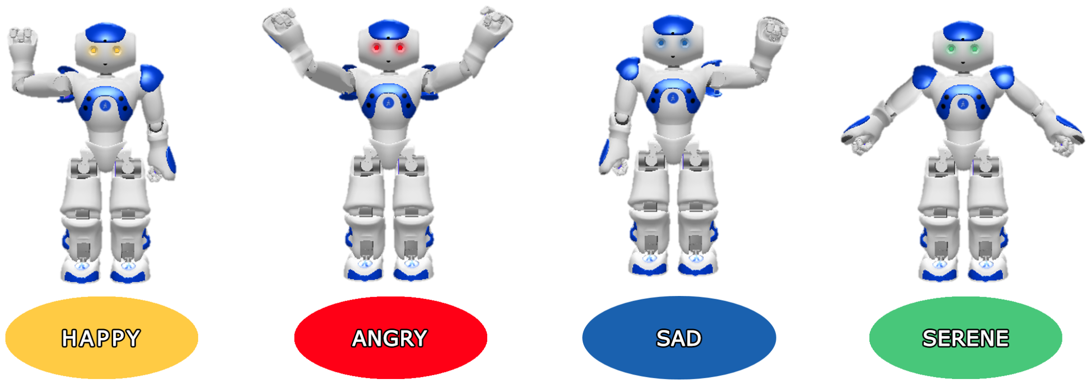


Figure 9.8: The four robot directional gestures.

Director module, and each of them is interpreted as one of the four songs that can be played in the next musical act. Each song belongs to an emotional category, which in turn is associated to a color. Thus, the robot will modify the color of his LEDs accordingly, to give a continuous reference to the music players even after the gesture movement is over.

	Inexpert	Quite Inexpert	Neutral	Expert	Very expert
1 Are you an computer expert?	23	28	33	22	12
	Never	**Rarely**	**Sometiomes**	**Usually**	**Often**
2 How frequently do you listen to symphonic music?	28	30	23	23	14

Table 9.2: Questionnaire items related to the predisposition torward robotics and synphonic music.

	1 Strongly disagree	2 Disagree	3 Neutral	4 Agree	5 Strongly Agree
4 Did the robot correctly fulfill the task?	15	2	28	46	22
5 Do you believe that the presence of a person instead of the robot might change the effectiveness of the concert?	21	25	33	21	18
6 Before the concert, did you expect an emotional involvement from the performance?	6	16	27	54	15
7 Did the robot meet your expectations?	11	19	28	37	23
8 Did you enjoy the concert?					

Table 9.3: The questionnaire filled from the users and the score attributed to each item.

9.4 Evaluation

After some preliminary tests, aimed at fixing bugs and tuning some parameters, the system has been tested during an actual concert, opened to an audience made by 118 spectators (63 males and 55 females). As explained before, each spectator was given with a mobile application, by means of which they were able to select a color during the concert. At the end of the concert, people filled a questionnaire consisting of questions to be answered with 5-points Likert scales.

The first two questions, reported in table 9.2 investigate user predisposition torward robotics and symphonic music.

The second part of the questionnaire was used to assess the audience perception during the concert. The details of questions and answers are reported in table 9.3.

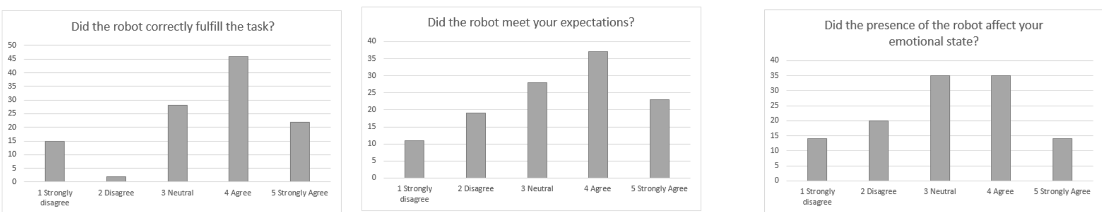


Figure 9.9: Users' answers to questions 4, 7 and 8.

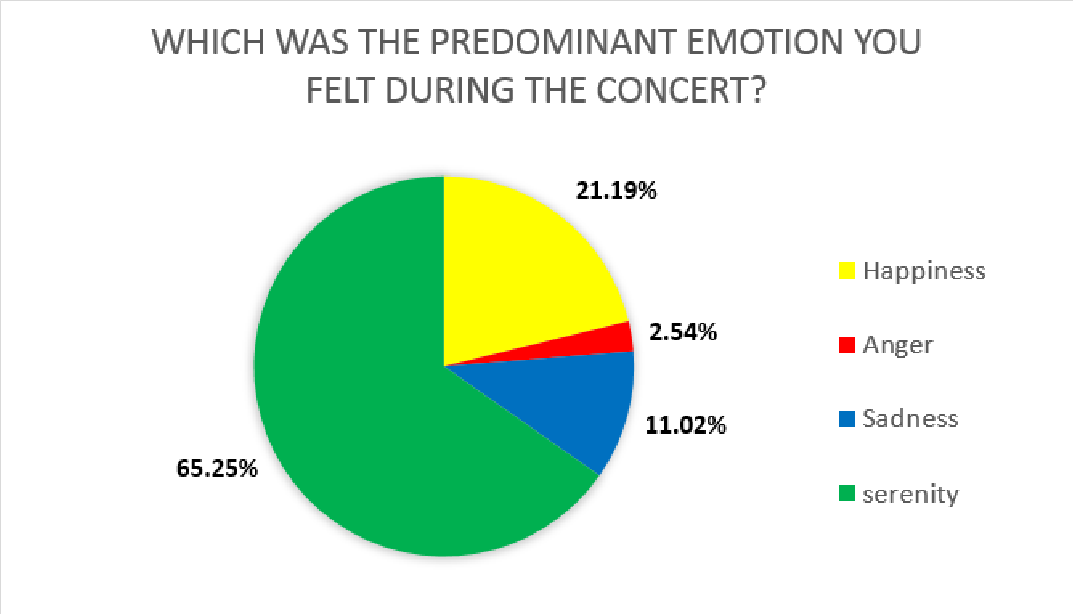


Figure 9.10: The prevalent audience emotions.

Finally, the questionnaire contained a question dedicated to the most perceived emotion, between the four proposed. As reported in figure 9.10 was serenity, with a predominance for the 65.25% of the audience.

The questionnaire was prepared to investigate users' attitude torward the performance and since it is not a standard questionnaire the Chronbach's values are reported: Cronbach Alpha 0.8279 Std. Alpha 0.8253 G6(smc) 0.8454 Average R 0.403. In table 9.3 are reported questions and answer provided from the audience.

Questions 4, 7 and 9 revealed interesting aspects of users' opinions. As shown in Figure 9.9, the majority of users enjoyed the concert, and no relations were noted between neither the level of expertise in informatics nor the users' habits in attending concerts. Moreover, most of the users felt like the robot correctly fulfill the task, and that its behavior meets users' expectations.

The observed differences in users' choices may be due to external factors, including the presence of the robot or in the absence of preliminary information and this point must be evaluated in a further study.

9.5 Conclusions

The hypothesis presented in this chapter is that is possible to realize a natural-like setting for HRI in which the humans and a robot coordinate behaviours and emotions [208] and communicate on the basis of shared phenomenal features in a common environment [209] to answer to the following *"Is it possible to convey humans' emotional response through the humanoid robot NAO?*

Many studies explored the emotional response in a Human Robot Interaction [19], [125] or [126]. The above mentioned study focused on a one-to-one interaction between a human and a robot, while in the study presented in this chapter I extend the investigation of emotional response over the whole audience, using the Nao robot for conveying these emotional response.

More in details, in this chapter it has been presented a system for conveying audience emotions to orchestra musicians through a humanoid robot, describing the architecture along with a brief overview of the algorithms used. First results collected after the deployment of the system in a real scenario have been discussed with particular attention to user's experience during and after the concert. From users' answers, we noted that most of the users enjoyed the concert, and the robot appeared to influence their opinions over the overall experience, wile this observation seems encouraging, further investigations are required to strengthen this hypothesis.

Finally robot was able to create a link between the audience and the orchestra, being the medium that conveyed peoples emotion during the overall concert experience. On the other hand, orchestra had the opportunity to know in real-time audience overall emotional state changing their performance accordingly to that state and creating an innovative and partecipative music experience for audience.

Chapter 10

Conclusions

The presented thesis aimed in providing an answer to the following research question: *is it possible to model human's responses during human robot interactions, using new measurable features?* To provide an answer, Human's feedbacks have been evaluated during Human Robot Interaction to derive measurable features to model human behaviour. The use of these features are the basis for the creation of humans' behavioural model with the final long time goal to make a robot aware of human's feedbacks.

My research fits into the field of Human-Robot Interaction using a multidisciplinary approach which took into account concepts and techniques borrowed from psychology, biology, cognitive science and medicine to better address the proposed research question.

Three paradigms of Human Robot Interaction have been taken into account:

1. Robot as Avatar.

2. Robot as Teammate.

3. Robot as Social Mediator.

Each paradigm has been explored in dedicated scenario measuring the features, as described in 10.1. Accordingly to **robot as Avatar** paradigm, the robot becomes an extension of the user himself and can be used to extend his physical presence, or accomplish action on his place.

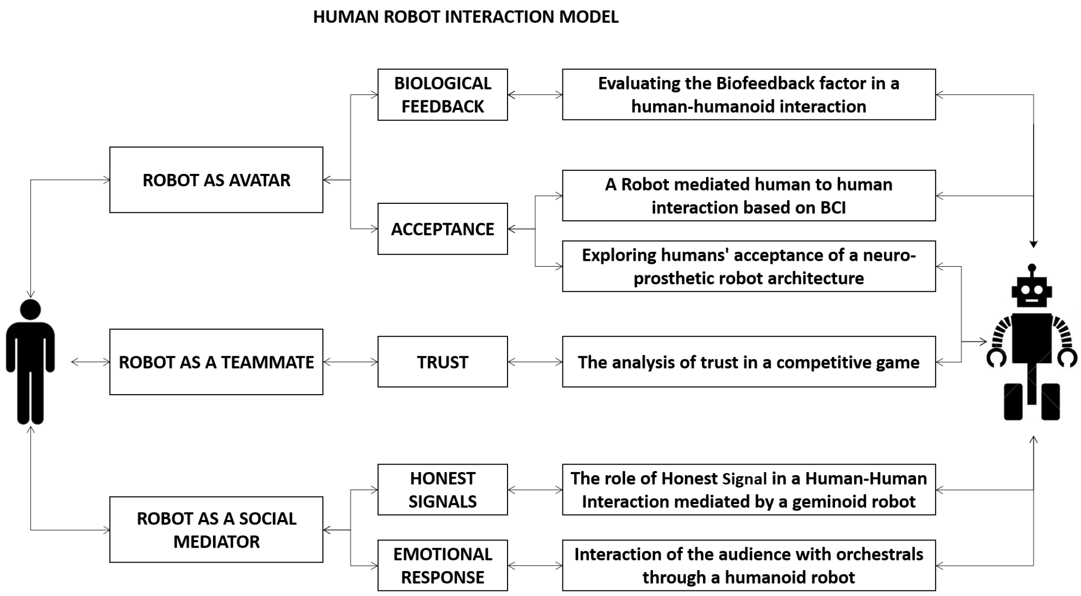


Figure 10.1: The proposed model.

The robot as Avatar paradigm have been evaluated using user's **Biological feedback** and **Acceptance**. First study, *Evaluating the Biofeedback factor in a human-humanoid interaction*, presented in chapter 4 evaluated the user's mental state during the interaction with a NAO robot used to fulfill user's needs.

This study validated a new paradigm to interact with a humanoid robot as avatar providing a positive answer to the question: *is the NAO Robot able to help a human in accomplishing his needs?*. This represents a further step in the studies on Brain Computer Interface where the output interface was typically represented by monitor [158], [69] or synthesized speech processor [159], rather then robots.

The study was conducted in collaboration with the Psychiatric department of the University of Palermo with healthy controls and ALS patients. ALS participants have shown a stronger **Biofeedback** (81.20%) rather than healthy controls (76.77%).

This result suggests that both ALS and healthy controls achieved a good precision during the online Session but, during the Robotic Session, the ALS performed statistically better. This conclusion strength the hypothesis that robot is perceived as a positive reinforcement to enhance user's performances, with particular attention to locked-in subjects.

Moreover the use of the robot produced an higher biofeedback in both groups reflecting an improvement of participants' mental state. The conclusion that can be drawn is that robot is perceived as a positive feedback for both healthy and ALS participants in the use of BCI.

Users' **Acceptance** has been evaluated in chapters 5 and 6. In Chapter 5, *a Robot mediated human to human interaction based on BCI*, the Brain Computer Interface system enabled ALS patient to communicate using a symbolic interface.

This study is novel in literature, since studies were mainly conducted to investigate BCI bit/rate (defined as time to spell a word) [170], system precision [171] and user's experience [172]. Moreover, the use of robot in a BCI based communication setting, has been quite neglected since robots were mainly used in controls scenarios [158].

From the results, it is possible to draw the conclusion that the use of a teleoperated robot helped in improving the sense of presence with the user and if a monitor or another electronic device would be used instead of robot, the interaction would not have been the same. Is therefore possible to conclude that the use of the robot provided an embodiment which helped to maintain a sense of presence.

The second study on user's acceptance is presented in chapter 6 *Exploring humans' acceptance of a neuro-prosthetic Robot architecture.* In this study a KUKA Robot has been brain controlled by an artist, as a neuro-prosthesis to create large artworks during Ars Electronica Festival 2017.

This study investigated the research question, *is it possible to model human's responses during human robot interactions, using new measurable features?* measuring the acceptance of a human-robot interaction in a real-world scenario over a wide sample.

This study is linked to the emergent question of the scientific community to provide assessment on robots' acceptance [183], [184]. Although acceptance of teleoperated robot has been widely explored, [25], [185], [186], the study presented in this chapter push forward the current state of the art since there are no prior examples of BCI-based control of an industrial robot used for art creation.

The results of the study, provided some interesting insights on people's current knowledge and predisposition torward robotics. In particular, industrial robots like KUKA and Brain Computer Interface are still not well known by the majority of people. Nevertheless people appear to have a positive attitude torward them, accepting them has useful technology in general and in particular to make art.

Results were evaluated from visitors interview in terms of technology knowledge, attitude, interaction and satisfaction. Presented results show that industrial robot like KUKA and Brain Computer Interface are still not well known by the majority of people. Nevertheless people appear to have a positive attitude torward them, accepting them has useful technology in general and in particular to make art. The performance has been considered innovative by the majority of people.

The Robot as Teammate has been explored in the study presented in chapter 7, *the analysis of trust in a BCI based Human-Humanoid Interaction*, where a human player played a competitive game against a bargain robot to evaluate human's mental model of trust during the interaction with the robot.

The goal of the study was to give an answer to the research question *is it possible to derive a model of human's trust from human's brain features during the interaction with a robot?* My study starts from the assumptions of [190] and [191] who used the social dynamic of cheating to investigate trust and robot agency. The main limitation of the above mentioned researches is represented by the use of speech and gestures to investigate the perception of the robot by the user. The research presented in this chapter 7 pushed forward this limitation introducing an additional measurement of user's trust: BCI features representing a model based on Heider's researches on trust [187].

Results demonstrated that it is possible to find a relation between biological features and **Trust**. In particular,the modification of the interaction is reflected by the BCI features. In addiction, from the obtained results is it possible to notice that the content of the interaction modify also the perception of the robot. In fact, the robot who played accordingly the rules or, even worse, who breaks the rules against its own interests is perceived as a broken tool subjected to system malfunction and therefore it was not not perceived as a trustable teammate.

Robot as a **social mediator** paradigm has been explored using different features: **honest signals** and **emotional response**. **Honest Signals** have been explored in chapter 8. Honest signalling have been considered as a subset of observable behaviour selected by a stable and equilibrium state of sharing information as the solution to coordination problems.

The study starts from Pentland's honest signal theory [122] that I extend to apply it to the research field of Human Robot Interaction. In fact Pentland's study was mainly focused on creating a sociometric instrument to measure human-human interaction [204] while my research goal is to apply honest signals theory to the human human interaction mediated by a humanoid robot.

Alhotugh due to the small sample, no definitive conclusions can be drawn, The study suggested two interesting results: First assessed that using the Geminoid robot permit to obtain better results during cooperative task between two users who never meet each other. Second, suggested that the Geminoid robot induced in a human a number of mimicrys comparable to those generated during interaction with another human being. The study therefore aims to be the starting point for future research on the interconnection between honest signals and humanoid robots.

The **Emotional Response** has been evaluated in chapter 9. In this study, users were invited to modify the ongoing of a live orchestral performance, held in collaboration with the conservatory of Palermo, using a mobile app as input and the NAO Robot as output of their overall emotional states.

The emotional response in a Human Robot Interaction has been investigated in several previous study [19], [125] or [126]. The main limitation of above mentioned studies is represented by the a one-to-one interaction between a human and a robot. I extended the above mentioned study investigating the emotional response over the whole audience and using the NAO robot for conveying the emotional response.

From the questionnaire submitted at the end of the performance, it emerged the majority of users enjoyed the concert, and no relations were noted between neither the level of expertise in informatics nor the users' habits in attending to such kind of concerts.

Moreover, most of the users perceived the robot correctly fulfilled the task, and that its behavior met users' expectations. The evaluation of the emotions perceived during the overall concert has shown that the most perceived emotion was serenity. Furthermore, the analysis of users' responses in terms of the chosen colors has shown a big variance in the variance of emotion along all the performance, with variation over Russel ideal circumplex. Interesting to notice that a relation between played pieces and emotions were found by the majority of people attending the concert.

In conclusion, the exploration of biological and perceptual feedback during a human-robot interaction within different paradigms (*robot as avatar, robot as teammate, robot as social mediator*) provided interesting results locating new features to measure and assess the interaction between a human and a robot.

In fact, from the results obtained for each study it is possible to derive features to measure humans' feedback. The further evolution of this research will be to equip a robot with a system based on these features to make the robot able to decode human beings mental and perceptual state and to react accordingly to their states and emotion based on analyzed features.

In general, the presented thesis represents a step forward in the current state of the art of Human Robot Interaction. For this reason, it becomes essential to assess the results achieved thus far in relation with well established features explored in Human Robot Interaction exploring, for instance the role on natural language, visual expressions and gestures to create a complete multimodal system able to better understand the extremely complex dimension of Human Robot Interaction.

The final long term goal will be the creation of a robotic architecture, based on cognitive and biological features, able to determine and to predict human's perceptive and biological feedback to realize a true and useful interaction for humans.

The creation of this robotic architecture will increase the naturalness and sociability of the interaction, moving the latter on the next stage: a stable and long term Human Robot Interaction to include robots in human daily-life.

Bibliography

[1] Michael A Goodrich and Alan C Schultz. Human-robot interaction: a survey. *Foundations and trends in human-computer interaction*, 1(3):203–275, 2007.

[2] Rachid Alami, Alin Albu-Schäffer, A. Bicchi, Rainer Bischoff, Raja Chatila, Alessandro De Luca, Agostino De Santis, Georges Giralt, Jérémie Guiochet, Gerd Hirzinger, et al. Safe and dependable physical human-robot interaction in anthropic domains: State of the art and challenges. In *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*, pages 1–16. IEEE, 2006.

[3] Thomas B Sheridan. Human–robot interaction: status and challenges. *Human factors*, 58(4):525–532, 2016.

[4] Y. Shahriari, E.W. Sellers, L.M. McCane, T.M. Vaughan, and D.J. Krusienski. Directional brain functional interaction analysis in patients with amyotrophic lateral sclerosis. In *Neural Engineering (NER), 2015 7th International IEEE/EMBS Conference on*, pages 972–975, 2015.

[5] Jintang Huang, Zhongmin Zhang, and Shaoyong Wang. Efficacy of the da vinci surgical system in colorectal surgery comparing with traditional laparoscopic surgery or open surgery: A meta-analysis. *International Journal of Advanced Robotic Systems*, 13(5):1729881416664849, 2016.

[6] Bishoy Morris. Robotic surgery: applications, limitations, and impact on surgical education. *Medscape General Medicine*, 7(3):72, 2005.

[7] Saeed Asadi Bagloee, Madjid Tavana, Mohsen Asadi, and Tracey Oliver. Autonomous vehicles: challenges, opportunities, and future implications for transportation policies. *Journal of Modern Transportation*, 24(4):284–303, 2016.

[8] Torsten Kammer, Philipp Brauner, Thiemo Leonhardt, and Ulrik Schroeder. Simulating lego mindstorms robots to facilitate teaching computer programming to school students. *Towards ubiquitous learning*, pages 196–209, 2011.

[9] Rolf Dieter Schraft, Birgit Graf, Andreas Traub, and Dirk John. A mobile robot platform for assistance and entertainment. *Industrial Robot: An International Journal*, 28(1):29–35, 2001.

[10] R. Sorbello, A. Chella, Marcello Giardina, Shuichi Nishio, and Hiroshi Ishiguro. An architecture for telenoid robot as empathic conversational android an architecture for telenoid robot as empathic conversational android companion for elderly people. *Intelligent Autonomous Systems 13, Advances in Intelligent Systems and Computing*, (302), 2015.

[11] S.Maria Anzalone, Elodie Tilmont, Sofiane Boucenna, Jean Xavier, Anne-Lise Jouen, Nicolas Bodeau, Koushik Maharatna, Mohamed Chetouani, David Cohen, MICHELANGELO Study Group, et al. How children with autism spectrum disorder behave and explore the 4-dimensional (spatial 3d+ time) environment during a joint attention induction task with a robot. *Research in Autism Spectrum Disorders*, 8(7):814–826, 2014.

[12] R. Tramonte, S.and Sorbello, Marcello Giardina, and A. Chella. Unipabci a novel general software framework for brain computer interface. In *Conference on Complex, Intelligent, and Software Intensive Systems*, pages 336–348. Springer, 2017.

[13] Anthony Giddens. Sociology (revised and updated with philip w. sutton). *Polity, Cambridge*, 2009.

[14] Dedre Gentner and Albert L Stevens. *Mental models.* Psychology Press, 2014.

[15] Christopher D Wickens, Justin G Hollands, Simon Banbury, and Raja Parasuraman. *Engineering psychology & human performance.* Psychology Press, 2015.

[16] Alex Mesoudi, Andrew Whiten, and Kevin N Laland. Towards a unified science of cultural evolution. *Behavioral and Brain Sciences*, 29(4):329–347, 2006.

[17] Cynthia Breazeal, Daphna Buchsbaum, Jesse Gray, David Gatenby, and Bruce Blumberg. Learning from and about others: Towards using imitation to bootstrap the social understanding of others by robots. *Artificial life*, 11(1-2):31–62, 2005.

[18] Brian Scassellati. Theory of mind for a humanoid robot. *Autonomous Robots*, 12(1):13–24, 2002.

[19] C. Breazeal. Social interactions in hri: the robot view. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 34(2):181–186, 2004.

[20] Cynthia Breazeal, Andrew Brooks, David Chilongo, Jesse Gray, Guy Hoffman, Cory Kidd, Hans Lee, Jeff Lieberman, and Andrea Lockerd. Working collaboratively with humanoid robots. In *Humanoid Robots, 2004 4th IEEE/RAS International Conference on*, volume 1, pages 253–272. IEEE, 2004.

[21] Jose M Carmena. Advances in neuroprosthetic learning and control. *PLoS biology*, 11(5):e1001561, 2013.

[22] Jennifer L Collinger, Stephen Foldes, Tim M Bruns, Brian Wodlinger, Robert Gaunt, and Douglas J Weber. Neuroprosthetic technology for individuals with spinal cord injury. *The journal of spinal cord medicine*, 36(4):258–272, 2013.

[23] Pratik Chhatbar. The future of implantable neuroprosthetic devices: ethical considerations. *Journal of long-term effects of medical implants*, 19(2), 2009.

[24] Kazuaki Tanaka, Hideyuki Nakanishi, and Hiroshi Ishiguro. Comparing video, avatar, and robot mediated communication: pros and cons of embodiment. In *International Conference on Collaboration Technologies*, pages 96–110. Springer, 2014.

[25] R. Sorbello, A. Chella, Carmelo Calí, Marcello Giardina, Shuichi Nishio, and Hiroshi Ishiguro. Telenoid android robot as an embodied perceptual social regulation medium engaging natural human–humanoid interaction. *Robotics and Autonomous Systems*, 62(9):1329–1341, 2014.

[26] Guy Hoffman and Cynthia Breazeal. Collaboration in human-robot teams. In *AIAA 1st Intelligent Systems Technical Conference*, page 6434, 2004.

[27] Kerstin Dautenhahn. Socially intelligent robots: dimensions of human–robot interaction. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 362(1480):679–704, 2007.

[28] Frank Hegel, Claudia Muhl, Britta Wrede, Martina Hielscher-Fastabend, and Gerhard Sagerer. Understanding social robots. In *Advances in Computer-Human Interactions, 2009. ACHI'09. Second International Conferences on*, pages 169–174. IEEE, 2009.

[29] Marcello Giardina, Salvatore Tramonte, Vito Gentile, Samuele Vinanzi, A. Chella, Salvatore Sorce, and R. Sorbello. *Conveying Audience Emotions Through Humanoid Robot Gestures to an Orchestra During a Live Musical Exhibition*, pages 249–261. Springer International Publishing, Advances in Intelligent Systems and Computing Volume 611, July 2017.

[30] T. Kanda, H. Ishiguro, T. Ono, M. Imai, and R. Nakatsu. Development and evaluation of an interactive humanoid robot robovie. In *Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on*, volume 2, pages 1848–1855. IEEE, 2002.

[31] Kerstin Dautenhahn, Chrystopher L Nehaniv, Michael L Walters, Ben Robins, Hatice Kose-Bagci, N Assif Mirza, and Mike Blow. Kaspar–a minimally expressive humanoid robot for human–robot interaction research. *Applied Bionics and Biomechanics*, 6(3-4):369–397, 2009.

[32] John-John Cabibihan, Hifza Javed, Marcelo Ang, and Sharifah Mariam Aljunied. Why robots? a survey on the roles and benefits of social robots in the therapy of children with autism. *International journal of social robotics*, 5(4):593–618, 2013.

[33] Cynthia Breazeal. Social robots for health applications. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 5368–5371. IEEE, 2011.

[34] Telenoid overview. `http://www.geminoid.jp/projects/kibans/Telenoid-overview.html`.

[35] Kohei Ogawa, Shuichi Nishio, Kensuke Koda, Koichi Taura, Takashi Minato, Carlos Toshinori Ishii, and Hiroshi Ishiguro. Telenoid: Tele-presence android for communication. In *ACM SIGGRAPH 2011 Emerging Technologies*, page 15. ACM, 2011.

[36] Ben Shneiderman. *Designing the user interface: strategies for effective human-computer interaction*. Pearson Education India, 2010.

[37] Helen Sharp, Yvonne Rogers, and Jenny Preece. Interaction design: beyond human-computer interaction. 2007.

[38] Jill Drury, Laurel D Riek, Alan D Christiansen, Zachary T Eyler-Walker, Andrea J Maggi, and David B Smith. Command and control of robot teams. In *Proc. of the Association of Unmanned Vehicles International (AUVSI) Conf*, 2003.

[39] Cheng Guo and Ehud Sharlin. Exploring the use of tangible user interfaces for human-robot interaction: a comparative study. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 121–130. ACM, 2008.

[40] Pramila Rani, Changchun Liu, Nilanjan Sarkar, and Eric Vanman. An empirical study of machine learning techniques for affect recognition in human–robot interaction. *Pattern Analysis and Applications*, 9(1):58–69, 2006.

[41] Phoebe Sengers and Bill Gaver. Staying open to interpretation: engaging multiple meanings in design and evaluation. In *Proceedings of the 6th conference on Designing Interactive systems*, pages 99–108. ACM, 2006.

[42] Juliet Corbin, Anselm Strauss, et al. Basics of qualitative research: Techniques and procedures for developing grounded theory. *Thousand Oaks*, 2008.

[43] Bernhard Graimann, Brendan Allison, and Gert Pfurtscheller. Brain–computer interfaces: A gentle introduction. In *Brain-Computer Interfaces*, pages 1–27. Springer, 2010.

[44] Jonathan Wolpaw and Elizabeth Winter Wolpaw. *Brain-computer interfaces: principles and practice*. OUP USA, 2012.

[45] Jonathan R Wolpaw and Dennis J McFarland. Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 101(51):17849–17854, 2004.

[46] José del R Millán, Pierre W Ferrez, Ferran Galán, Eileen Lew, and Ricardo Chavarriaga. Non-invasive brain-actuated interaction. In *Advances in Brain, Vision, and Artificial Intelligence*, pages 438–447. Springer, 2007.

[47] Michal Teplan. Fundamentals of eeg measurement. *Measurement science review*, 2(2):1–11, 2002.

[48] Amjed S Al-Fahoum and Ausilah A Al-Fraihat. Methods of eeg signal features extraction using linear analysis in frequency and time-frequency domains. *ISRN neuroscience*, 2014, 2014.

[49] BS Oken, MC Salinsky, and SM Elsas. Vigilance, alertness, or sustained attention: physiological basis and measurement. *Clinical neurophysiology*, 117(9):1885–1901, 2006.

[50] Brian H Bland. The physiology and pharmacology of hippocampal formation theta rhythms. *Progress in neurobiology*, 26(1):1–54, 1986.

[51] Brian H Bland, Scott D Oddie, and Luis V Colom. Mechanisms of neural synchrony in the septohippocampal pathways underlying hippocampal theta generation. *Journal of Neuroscience*, 19(8):3223–3237, 1999.

[52] J Gallinat, D Kunz, D Senkowski, T Kienast, F Seifert, F Schubert, and A Heinz. Hippocampal glutamate concentration predicts cerebral theta oscillations during cognitive processing. *Psychopharmacology*, 187(1):103–111, 2006.

[53] Adam MP Miller, Brendan J Frick, David M Smith, Jelena Radulovic, and Kevin A Corcoran. Network oscillatory activity driven by context memory processing is differently regulated by glutamatergic and cholinergic neurotransmission. *Neurobiology of Learning and Memory*, 2017.

[54] Erik W Schomburg, A. Fernández-Ruiz, Kenji Mizuseki, Antal Berényi, Costas A Anastassiou, Christof Koch, and György Buzsáki. Theta phase segregation of input-specific gamma patterns in entorhinal-hippocampal networks. *Neuron*, 84(2):470–485, 2014.

[55] Stuart N Baker. Oscillatory interactions between sensorimotor cortex and the periphery. *Current opinion in neurobiology*, 17(6):649–655, 2007.

[56] Elodie Lalo, Thomas Gilbertson, Louise Doyle, Vincenzo Di Lazzaro, Beatrice Cioni, and Peter Brown. Phasic increases in cortical beta activity are associated with alterations in sensory processing in the human. *Experimental brain research*, 177(1):137–145, 2007.

[57] Xiaoxuan Jia and Adam Kohn. Gamma rhythms in the brain. *PLoS biology*, 9(4):e1001045, 2011.

[58] Edgar Douglas Adrian. Olfactory reactions in the brain of the hedgehog. *The Journal of physiology*, 100(4):459–473, 1942.

[59] J Andrew Henrie and Robert Shapley. Lfp power spectra in v1 cortex: the graded effect of stimulus contrast. *Journal of neurophysiology*, 94(1):479–490, 2005.

[60] 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.

[61] Pascal Fries, John H Reynolds, Alan E Rorie, and Robert Desimone. Modulation of oscillatory neuronal synchronization by selective visual attention. *Science*, 291(5508):1560–1563, 2001.

[62] Philipp Berens, Georgios A Keliris, Alexander S Ecker, Nikos K Logothetis, and Andreas S Tolias. Feature selectivity of the gamma-band of the local field potential in primate primary visual cortex. *Frontiers in neuroscience*, 2(2):199, 2008.

[63] Jing Liu and William T Newsome. Local field potential in cortical area mt: stimulus tuning and behavioral correlations. *Journal of Neuroscience*, 26(30):7779–7790, 2006.

[64] Michael J Kahana. The cognitive correlates of human brain oscillations. *Journal of Neuroscience*, 26(6):1669–1672, 2006.

[65] 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.

[66] Arnaud Delorme, Makoto Miyakoshi, Tzyy-Ping Jung, and Scott Makeig. Grand average erp-image plotting and statistics: A method for comparing variability in event-related single-trial eeg activities across subjects and conditions. *Journal of neuroscience methods*, 250:3–6, 2015.

[67] John Polich and José R Criado. Neuropsychology and neuropharmacology of p3a and p3b. *International Journal of Psychophysiology*, 60(2):172–185, 2006.

[68] John Polich. Updating p300: an integrative theory of p3a and p3b. *Clinical neurophysiology*, 118(10):2128–2148, 2007.

[69] Emanuel Donchin and Michael GH Coles. Is the p300 component a manifestation of context updating? *Behavioral and brain sciences*, 11(03):357–374, 1988.

[70] DC Dugdale, DB Hoch, and D Zieve. Amyotrophic lateral sclerosis. *ADAM Medical Encyclopedia*, 2010.

[71] LA. Farwell and E. 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.

[72] Dean J Krusienski, Eric W Sellers, François Cabestaing, Sabri Bayoudh, 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.

[73] Hilit Serby, Elad Yom-Tov, and Gideon F Inbar. An improved p300-based brain-computer interface. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 13(1):89–98, 2005.

[74] 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.

[75] Ronald Fisher. Statistical methods and scientific induction. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 69–78, 1955.

[76] Eric W Sellers, Dean J Krusienski, Dennis J McFarland, Theresa M Vaughan, and Jonathan R Wolpaw. A p300 event-related potential brain–computer interface (bci): the effects of matrix size and inter stimulus interval on performance. *Biological psychology*, 73(3):242–252, 2006.

[77] Alain Rakotomamonjy and Vincent Guigue. Bci competition iii: dataset ii-ensemble of svms for bci p300 speller. *Biomedical Engineering, IEEE Transactions on*, 55(3):1147–1154, 2008.

[78] Jose Miguel Leiva-Murillo and A. Artés-Rodríguez. Maximization of mutual information for supervised linear feature extraction. *Neural Networks, IEEE Transactions on*, 18(5):1433–1441, 2007.

[79] George Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of control, signals and systems*, 2(4):303–314, 1989.

[80] Nikolaj Blom, Steen Gammeltoft, and Søren Brunak. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *Journal of molecular biology*, 294(5):1351–1362, 1999.

[81] Yuanqing Li, Jiahui Pan, Fei Wang, and Zhuliang Yu. A hybrid bci system combining p300 and ssvep and its application to wheelchair control. *Biomedical Engineering, IEEE Transactions on*, 60(11):3156–3166, 2013.

[82] Gert Pfurtscheller, Brendan Z Allison, Günther Bauernfeind, Clemens Brunner, Teodoro Solis Escalante, Reinhold Scherer, Thorsten O Zander, Gernot Mueller-Putz, Christa Neuper, and Niels Birbaumer. The hybrid bci. *Frontiers in neuroscience*, 4:3, 2010.

[83] Darren O'Doherty, Yogesh Kumar Meena, Haider Raza, Hubert Cecotti, and Girijesh Prasad. Exploring gaze-motor imagery hybrid brain-computer interface design. In *Bioinformatics and Biomedicine (BIBM), 2014 IEEE International Conference on*, pages 335–339. IEEE, 2014.

[84] Yogesh Meena, Girijesh Prasad, Hubert Cecotti, and KongFatt Wong-Lin. Simultaneous gaze and motor imagery hybrid bci increases single-trial detection performance: a compatible incompatible study. IEEE EMBS, 2015.

[85] Kazuo Tanaka, Kazuyuki Matsunaga, and Hua O Wang. Electroencephalogram-based control of an electric wheelchair. *IEEE transactions on robotics*, 21(4):762–766, 2005.

[86] Moore Vora J.Y., Allison B.Z. One relatively early project presented a p300 bci to direct a robotic arm to make coffee. *. Society for Neuroscience Abstract, 30, Program No. 421.19.*, 2004.

[87] Nicholas Waytowich, Andrew Henderson, Dean Krusienski, and Daniel Cox. Robot application of a brain computer interface to staubli tx40 robots-early stages. In *World Automation Congress (WAC), 2010*, pages 1–6. IEEE, 2010.

[88] Christian J Bell, Pradeep Shenoy, Rawichote Chalodhorn, and Rajesh PN Rao. Control of a humanoid robot by a noninvasive brain–computer interface in humans. *Journal of neural engineering*, 5(2):214, 2008.

[89] A. Chella, Enrico Pagello, Emanuele Menegatti, R. Sorbello, S.Maria Anzalone, Francesco Cinquegrani, Luca Tonin, Francesco Piccione, K Prifitis, Claudia Blanda, et al. A bci teleoperated museum robotic guide. In *Complex, Intelligent and Software Intensive Systems, 2009. CISIS'09. International Conference on*, pages 783–788. IEEE, 2009.

[90] Maryam Alimardani, Shuichi Nishio, and Hiroshi Ishiguro. The importance of visual feedback design in bcis; from embodiment to motor imagery learning. *PloS one*, 11(9):e0161945, 2016.

[91] Emmanuele Tidoni, Mohammad Abu-Alqumsan, Daniele Leonardis, Christoph Kapeller, Gabriele Fusco, Christoph Guger, Cristoph Hintermueller, Angelika Peer, A. Frisoli, Franco Tecchia, et al. Local and remote cooperation with virtual and robotic agents: a p300 bci study in healthy and people living with spinal cord injury. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2016.

[92] Torbjørn S Dahl and Maged N Kamel Boulos. Robots in health and social care: A complementary technology to home care and telehealthcare? *Robotics*, 3(1):1–21, 2013.

[93] Rupert Ortner, David Ram, Alexander Kollreider, Harald Pitsch, Joanna Wojtowicz, and Günter Edlinger. Human-computer confluence for rehabilitation purposes after stroke. In *Virtual, Augmented and Mixed Reality. Systems and Applications*, pages 74–82. Springer, 2013.

[94] Marco D Comerchero and John Polich. P3a and p3b from typical auditory and visual stimuli. *Clinical Neurophysiology*, 110(1):24–30, 1999.

[95] Rik van Dinteren, Martijn Arns, Marijtje LA Jongsma, and Roy PC Kessels. P300 development across the lifespan: a systematic review and meta-analysis. *PLoS One*, 9(2):e87347, 2014.

[96] Robert GD Steel and H James. Principles and procedures of statistics: with special reference to the biological sciences. Technical report, New York, US: McGraw-Hill, 1960.

[97] Joseph N Mak, Dennis J McFarland, Theresa M Vaughan, Lynn M McCane, Phillippa Z Tsui, Debra J Zeitlin, Eric W Sellers, and Jonathan R Wolpaw. Eeg correlates of p300-based brain computer interface (bci) performance in people with amyotrophic lateral sclerosis. *Journal of neural engineering*, 9(2):026014, 2012.

[98] Yuanqing Li, Jiahui Pan, Yanbin He, Fei Wang, Steven Laureys, Qiuyou Xie, and Ronghao Yu. Detecting number processing and mental calculation in patients with disorders of consciousness using a hybrid brain-computer interface system. *BMC neurology*, 15(1):1, 2015.

[99] S. Reisman. Measurement of physiological stress. In *Proc. IEEE 1997 23rd Northeast Bioengineering Conf.*, pages 21–23, May 1997.

[100] N. Sulaiman, M. N. Taib, S. Lias, Z. H. Murat, S. A. Mohd Aris, M. Mustafa, and N. A. Rashid. Development of EEG-based stress index. In *Proc. Int Biomedical Engineering (ICoBE) Conf*, pages 461–466, February 2012.

[101] Y Tran, RA Thuraisingham, N Wijesuriya, HT Nguyen, and A Craig. Detecting neural changes during stress and fatigue effectively: a comparison of spectral analysis and sample entropy. In *3rd International IEEE/EMBS Conference on Neural Engineering, 2007. CNE'07*, pages 350–353. IEEE, 2007.

[102] Yichuan Liu, Hasan Ayaz, Banu Onaral, and Patricia A Shewokis. Eeg band powers for characterizing user engagement in p300-bci. In *Neural Engineering (NER), 2013 6th International IEEE/EMBS Conference on*, pages 1066–1069. IEEE, 2013.

[103] Saroj KL Lal, Ashley Craig, Peter Boord, Les Kirkup, and Hung Nguyen. Development of an algorithm for an eeg-based driver fatigue countermeasure. *Journal of Safety Research*, 34(3):321–328, 2003.

[104] Christos Papadelis, Zhe Chen, Chrysoula Kourtidou-Papadeli, Panagiotis D Bamidis, Ioanna Chouvarda, Evangelos Bekiaris, and Nikos Maglaveras. Monitoring sleepiness with on-board electrophysiological recordings for preventing sleep-deprived traffic accidents. *Clinical Neurophysiology*, 118(9):1906–1922, 2007.

[105] Zhenhu Liang, Yinghua Wang, Xue Sun, Duan Li, Logan J Voss, Jamie W Sleigh, Satoshi Hagihira, and Xiaoli Li. Eeg entropy measures in anesthesia. *Frontiers in computational neuroscience*, 9:16, 2015.

[106] D Harrison McKnight, Vivek Choudhury, and Charles Kacmar. The impact of initial consumer trust on intentions to transact with a web site: a trust building model. *The Journal of Strategic Information Systems*, 11(3):297–323, 2002.

[107] Michael Reid and Yair Levy. Integrating trust and computer self-efficacy with tam: An empirical assessment of customers' acceptance of banking information systems (bis) in jamaica. *Journal of Internet Banking and Commerce*, 12(3):1–17, 2008.

[108] David A Scheinberg, Carlos H Villa, Freddy E Escorcia, and Michael R McDevitt. Conscripts of the infinite armada: systemic cancer therapy using nanomaterials. *Nature Reviews Clinical Oncology*, 7(5):266–276, 2010.

[109] George Charalambous, Sarah Fletcher, and Philip Webb. The development of a scale to evaluate trust in industrial human-robot collaboration. *International Journal of Social Robotics*, 8(2):193–209, 2016.

[110] John D Lee and Katrina A See. Trust in automation: Designing for appropriate reliance. *Human factors*, 46(1):50–80, 2004.

[111] Ewart J de Visser, Marvin Cohen, Amos Freedy, and Raja Parasuraman. A design methodology for trust cue calibration in cognitive agents. In *International Conference on Virtual, Augmented and Mixed Reality*, pages 251–262. Springer, 2014.

[112] Munjal Desai, Kristen Stubbs, Aaron Steinfeld, and Holly Yanco. Creating trustworthy robots: Lessons and inspirations from automated systems. 2009.

[113] S.M Anzalone, Sofiane Boucenna, Serena Ivaldi, and Mohamed Chetouani. Evaluating the engagement with social robots. *International Journal of Social Robotics*, 7(4):465–478, 2015.

[114] Serena Ivaldi, S.M Anzalone, Woody Rousseau, Olivier Sigaud, and Mohamed Chetouani. Robot initiative in a team learning task increases the rhythm of interaction but not the perceived engagement. *Frontiers in neurorobotics*, 8, 2014.

[115] Cynthia Breazeal. Emotion and sociable humanoid robots. *International Journal of Human-Computer Studies*, 59(1):119–155, 2003.

[116] Victoria Groom and Clifford Nass. Can robots be teammates?: Benchmarks in human–robot teams. *Interaction Studies*, 8(3):483–500, 2007.

[117] M. Mori. The uncanny valley. *Energy*, 7(4):33–35, 1970.

[118] Frank MF Verberne, Jaap Ham, and Cees JH Midden. Trusting a virtual driver that looks, acts, and thinks like you. *Human factors*, 57(5):895–909, 2015.

[119] Amos Freedy, Ewart DeVisser, Gershon Weltman, and Nicole Coeyman. Measurement of trust in human-robot collaboration. In *Collaborative Technologies and Systems, 2007. CTS 2007. International Symposium on*, pages 106–114. IEEE, 2007.

[120] Faezeh Rahbar, S.M Anzalone, G. Varni, Elisabetta Zibetti, Serena Ivaldi, and Mohamed Chetouani. Predicting extraversion from non-verbal features during a face-to-face human-robot interaction. In *International Conference on Social Robotics*, pages 543–553. Springer, 2015.

[121] G. Varni, G. Volpe, and A. Camurri. A system for real-time multimodal analysis of non-verbal affective social interaction in user-centric media. *IEEE Transactions on Multimedia*, 12(6):576–590, 2010.

[122] Alex Pentland. Honest signals: How they shape our world (bradford books). 2008.

[123] Alex Sandy Pentland. Social signal processing [exploratory dsp]. *Signal Processing Magazine, IEEE*, 24(4):108–111, 2007.

[124] Daniel Olguín Olguín, Benjamin N Waber, Taemie Kim, Akshay Mohan, Koji Ara, and Alex Pentland. Sensible organizations: Technology and methodology for automatically measuring organizational behavior. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 39(1):43–55, 2009.

[125] Vito Gentile, Fabrizio Milazzo, A. Sorce, S.and Gentile, Giovanni Pilato, and Agnese Augello. Body Gestures and Spoken Sentences: a Novel Approach for Revealing User's Emotions. In *Proceedings of 11th International Conference on Semantic Computing (IEEE ICSC 2017)*, 2017.

[126] Sascha Meudt, Miriam Schmidt-Wack, Frank Honold, Felix Schüssel, Michael Weber, Friedhelm Schwenker, and Günther Palm. *Going Further in Affective Computing: How Emotion Recognition Can Improve Adaptive User Interaction*, pages 73–103. Springer International Publishing, Cham, 2016.

[127] Roddy Cowie and Randolph R. Cornelius. Describing the emotional states that are expressed in speech. *Speech Commun.*, 40(1-2):5–32, 2003.

[128] Paul Ekman. *Basic Emotions*, pages 45–60. John Wiley & Sons, Ltd, 2005.

[129] James A. Russell. A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6):1161–1178, 1980.

[130] Jonathan Posner, James A. Russel, and Bradley S. Peterson. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, 17(3):715–734, 2005.

[131] Agnese Augello, Ignazio Infantino, Giovanni Pilato, Riccardo Rizzo, and Filippo Vella. Binding representational spaces of colors and emotions for creativity. *Biologically Inspired Cognitive Architectures*, 5:64 – 71, 2013.

[132] Marko Tkalčič, Berardina De Carolis, Marco de Gemmis, Ante Odić, and Andrej Košir. *Introduction to Emotions and Personality in Personalized Systems*, pages 3–11. Springer International Publishing, Cham, 2016.

[133] Guy Hoffman, Shira Bauman, and Keinan Vanunu. Robotic experience companionship in music listening and video watching. *Personal and Ubiquitous Computing*, 20(1):51–63, 2016.

[134] Angelica Lim, Tetsuya Ogata, and Hiroshi G. Okuno. Towards expressive musical robots: a cross-modal framework for emotional gesture, voice and music. *EURASIP Journal on Audio, Speech, and Music Processing*, 2012(1):3, 2012.

[135] Louis McCallum and Peter W. McOwan. Face the music and glance: How nonverbal behaviour aids human robot relationships based in music. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, HRI '15, pages 237–244, New York, NY, USA, 2015. ACM.

[136] Birgitta Burger and Roberto Bresin. Communication of musical expression by means of mobile robot gestures. *Journal on Multimodal User Interfaces*, 3(1):109–118, 2010.

[137] L. Brown and A. M. Howard. Gestural behavioral implementation on a humanoid robotic platform for effective social interaction. In *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, pages 471–476, 2014.

[138] 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: teleoperators and virtual environments*, 19(1):35–53, 2010.

[139] Steven G Mason and Gary E Birch. A general framework for brain-computer interface design. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 11(1):70–85, 2003.

[140] Thilo Hinterberger, Jürgen Mellinger, and Niels Birbaumer. The thought translation device: Structure of a multimodal brain-computer communication system. In *Neural Engineering, 2003. Conference Proceedings. First International IEEE EMBS Conference on*, pages 603–606. IEEE, 2003.

[141] Thilo Hinterberger, Femke Nijboer, Andrea Kübler, Tamara Matuz, Adrian Furdea, Ursula Mochty, Miguel Jordan, Thomas Navin Lal, N Jeremy Hill, Jürgen Mellinger, et al. Brain-computer interfaces for communication in paralysis: A clinical experimental approach. *Towards Brain-Computer Interfacing.*, pages 43–64, 2007.

[142] Cecil Hastings Jr, Frederick Mosteller, John W Tukey, and Charles P Winsor. Low moments for small samples: a comparative study of order statistics. *The Annals of Mathematical Statistics*, pages 413–426, 1947.

[143] Tao Chen and Elaine Martin. Bayesian linear regression and variable selection for spectroscopic calibration. *Analytica chimica acta*, 631(1):13–21, 2009.

[144] Christopher M Bishop. *Pattern recognition and machine learning.* springer, 2006.

[145] David JC MacKay. Bayesian interpolation. *Neural computation*, 4(3):415–447, 1992.

[146] T Kaufmann, S M Schulz, C Grunzinger, and A Kubler. Flashing characters with famous faces improves erp-based brain computer interface performance. *Journal of Neural Engineering*, 8(5):056016, 2011.

[147] Michal Teplan et al. Fundamentals of eeg measurement. *Measurement science review*, 2(2):1–11, 2002.

[148] J. A. Wilson, J. Mellinger, G. Schalk, and J. Williams. A procedure for measuring latencies in brain x2013;computer interfaces. *IEEE Transactions on Biomedical Engineering*, 57(7):1785–1797, 2010.

[149] Rossella Spataro, A. Chella, Brendan Allison, Marcello Giardina, R. Sorbello, Christoph Tramonte, S.and Guger, and Vincenzo La Bella. Reaching and grasping a glass of water by locked-in als patients through a bci-controlled humanoid robot. *Frontiers in Human Neuroscience*, 11:68, 2017.

[150] R. Sorbello, Marcello Tramonte, S.and Giardina, Vincenzo La Bella, Rossella Spataro, Brendan Allison, Christoph Guger, and A. Chella. A human-humanoid interaction through the use of bci for locked-in als patients using neuro-biological feedback fusion. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2017.

[151] David MJ Tax and Robert PW Duin. Using two-class classifiers for multiclass classification. In *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, volume 2, pages 124–127. IEEE, 2002.

[152] Andrew Myrden and Tom Chau. Effects of user mental state on eeg-bci performance. *Frontiers in human neuroscience*, 9, 2015.

[153] Peter C Austin. A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003. *Statistics in medicine*, 27(12):2037–2049, 2008.

[154] Marshal F Folstein, Susan E Folstein, and Paul R McHugh. "mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research*, 12(3):189–198, 1975.

[155] Janet Grace and Paul F Malloy. *Frontal systems behavior scale: professional manual.* Psychological Assessment Resources, Incorporated, 2000.

[156] Eva Maria Hammer, Sonja Häcker, Martin Hautzinger, Thomas D Meyer, and Andrea Kübler. Validity of the als-depression-inventory (adi-12)-a new screening instrument for depressive disorders in patients with amyotrophic lateral sclerosis. *Journal of affective disorders*, 109(1):213–219, 2008.

[157] Aaron T Beck and Alice Beamesderfer. *Assessment of depression: the depression inventory.* Karger Publishers, 1974.

[158] Sarah N Abdulkader, Ayman Atia, and Mostafa-Sami M Mostafa. Brain computer interfacing: Applications and challenges. *Egyptian Informatics Journal*, 16(2):213–230, 2015.

[159] Sumit Soman and BK Murthy. Using brain computer interface for synthesized speech communication for the physically disabled. *Procedia Computer Science*, 46:292–298, 2015.

[160] Maja Matarić. Socially assistive robotics: human-robot interaction methods for creating robots that care. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, pages 333–333. ACM, 2014.

[161] Juan Fasola and Maja Mataric. A socially assistive robot exercise coach for the elderly. *Journal of Human-Robot Interaction*, 2(2):3–32, 2013.

[162] Helge Hüttenrauch, Kerstin Severinson Eklundh, Anders Green, and Elin A Topp. Investigating spatial relationships in human-robot interaction. In *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*, pages 5052–5059. IEEE, 2006.

[163] Elizabeth Broadbent, Rie Tamagawa, Anna Patience, Brett Knock, Ngaire Kerse, Karen Day, and Bruce A MacDonald. Attitudes towards health-care robots in a retirement village. *Australasian journal on ageing*, 31(2):115–120, 2012.

[164] Lewis P Rowland and Neil A Shneider. Amyotrophic lateral sclerosis. *New England Journal of Medicine*, 344(22):1688–1700, 2001.

[165] Miriam Kyselo and Ezequiel Di P. Locked-in syndrome: a challenge for embodied cognitive science. *Phenomenology and the Cognitive Sciences*, 14(3):517–542, 2015.

[166] Lynn M McCane, Susan M Heckman, and et al. McFarland. P300-based brain computer interface (bci) event-related potentials (erps): People with amyotrophic lateral sclerosis (als) vs age matched controls. *Clinical Neurophysiology*, 126(11):2124–2131, 2015.

[167] A Chio, G Logroscino, BJ Traynor, J Collins, JC Simeone, LA Goldstein, and LA White. Global epidemiology of amyotrophic lateral sclerosis: a systematic review of the published literature. *Neuroepidemiology*, 41(2):118–130, 2013.

[168] R. Tramonte, S.and Sorbello, Marcello Giardina, and A. Chella. Unipabci a novel general software framework for brain computer interface. In *Conference on Complex, Intelligent, and Software Intensive Systems*, pages 336–348. Springer, 2017.

[169] Robert I Jennrich and Paul Sampson. Stepwise discriminant analysis. *Statistical methods for digital computers*, 3:77–95, 1977.

[170] Faraz Akram, Seung Moo Han, and Tae-Seong Kim. An efficient word typing p300-bci system using a modified t9 interface and random forest classifier. *Computers in biology and medicine*, 56:30–36, 2015.

[171] Radhika Swarnkar, AG Keskar, PM Shyam Prasad, and NC Shivprakash. A new approach to detect p300 in a single trial based on pca and svm classifier. In *Region 10 Symposium (TENSYMP), 2016 IEEE*, pages 355–360. IEEE, 2016.

[172] Andrea Kübler, Elisa M Holz, Angela Riccio, Claudia Zickler, Tobias Kaufmann, Sonja C Kleih, Pit Staiger-Sälzer, Lorenzo Desideri, Evert-Jan Hoogerwerf, and Donatella Mattia. The user-centered design as novel perspective for evaluating the usability of bci-controlled applications. *PLoS One*, 9(12):e112392, 2014.

[173] Chet T Moritz, Steve I Perlmutter, and Eberhard E Fetz. Direct control of paralysed muscles by cortical neurons. *Nature*, 456(7222):639–642, 2008.

[174] Christian Ethier, Emily R Oby, MJ Bauman, and Lee E Miller. Restoration of grasp following paralysis through brain-controlled stimulation of muscles. *Nature*, 485(7398):368–371, 2012.

[175] Eric B Knudsen, Marissa E Powers, and Karen A Moxon. Dissociating movement from movement timing in the rat primary motor cortex. *Journal of Neuroscience*, 34(47):15576–15586, 2014.

[176] D. A. Todd, P. J. McCullagh, M. D. Mulvenna, and G. Lightbody. Investigating the use of brain-computer interaction to facilitate creativity. In *Proceedings of the 3rd Augmented Human International Conference*, AH '12, pages 19:1–19:8, New York, NY, USA, 2012. ACM.

[177] Jana I Münßinger, Sebastian Halder, Sonja C Kleih, Adrian Furdea, Valerio Raco, Adi Hösle, and Andrea Kübler. Brain painting: first evaluation of a new brain–computer interface application with als-patients and healthy volunteers. *Frontiers in neuroscience*, 4, 2010.

[178] Geoffrey Biggs and Bruce MacDonald. A survey of robot programming systems. In *Proceedings of the Australasian conference on robotics and automation*, pages 1–3, 2003.

[179] Filippo Sanfilippo, Lars Ivar Hatledal, Houxiang Zhang, Massimiliano Fago, and Kristin Y Pettersen. Controlling kuka industrial robots: Flexible communication interface jopenshowvar. *IEEE Robotics & Automation Magazine*, 22(4):96–109, 2015.

[180] Terence W Picton. The p300 wave of the human event-related potential. *Journal of clinical neurophysiology*, 9(4):456–479, 1992.

[181] Jennifer Y Bennington and John Polich. Comparison of p300 from passive and active tasks for auditory and visual stimuli. *International Journal of Psychophysiology*, 34(2):171–177, 1999.

[182] Christian Martens, Oliver Prenzel, and Axel Gräser. The rehabilitation robots friend-i & ii: Daily life independency through semi-autonomous task-execution. In *Rehabilitation robotics*. InTech, 2007.

[183] Jenay M Beer, Akanksha Prakash, Tracy L Mitzner, and Wendy A Rogers. Understanding robot acceptance. Technical report, Georgia Institute of Technology, 2011.

[184] Friederike Eyssel, Dieta Kuchenbrandt, Simon Bobinger, Laura de Ruiter, and Frank Hegel. If you sound like me, you must be more human': On the interplay of robot and user features on human-robot acceptance and anthropomorphism. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, pages 125–126. ACM, 2012.

[185] Ruth Maria Stock and Moritz Merkle. A service robot acceptance model: User acceptance of humanoid robots during service encounters. In *Pervasive Computing and Communications Workshops (PerCom Workshops), 2017 IEEE International Conference on*, pages 339–344. IEEE, 2017.

[186] Yan Wu, Wei Liang Chan, Yanan Li, Keng Peng Tee, Rui Yan, and Dilip K Limbu. Improving human-robot interactivity for tele-operated industrial and service robot applications. In *Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM), 2015 IEEE 7th International Conference on*, pages 153–158. IEEE, 2015.

[187] Fritz Heider. Interpersonal relations. *New YorN: Wiley*, 1958.

[188] Tatiana Sitnikova, Phillip J Holcomb, Kristi A Kiyonaga, and Gina R Kuperberg. Two neurocognitive mechanisms of semantic integration during the comprehension of visual real-world events. *Journal of cognitive neuroscience*, 20(11):2037–2057, 2008.

[189] Hideyuki Takahashi, Chinatsu Saito, Hiroyuki Okada, and Takashi Omori. An investigation of social factors related to online mentalizing in a human-robot competitive game. *Japanese Psychological Research*, 55(2):144–153, 2013.

[190] Elaine Short, Justin Hart, Michelle Vu, and Brian Scassellati. No fair!! an interaction with a cheating robot. In *Human-Robot Interaction (HRI), 2010 5th ACM/IEEE International Conference on*, pages 219–226. IEEE, 2010.

[191] Daniel Ullman, Lolanda Leite, Jonathan Phillips, Julia Kim-Cohen, and Brian Scassellati. Smart human, smarter robot: How cheating affects perceptions of social agency. In *Proceedings of the Cognitive Science Society*, volume 36, 2014.

[192] Fritz Heider and Marianne Simmel. An experimental study of apparent behavior. *The American journal of psychology*, 57(2):243–259, 1944.

[193] F. Heider. *The psychology of interpersonal relations.* Lawrence Erlbaum, 1982.

[194] Alex Pentland and Tracy Heibeck. *Honest signals.* MIT press Cambridge, MA, 2008.

[195] Raoul Bell, Julia Sasse, Malte Möller, Daniela Czernochowski, Susanne Mayr, and Axel Buchner. Event-related potentials in response to cheating and cooperation in a social dilemma game. *Psychophysiology*, 53(2):216–228, 2016.

[196] Marta Kutas and Kara D Federmeier. Thirty years and counting: finding meaning in the n400 component of the event-related brain potential (erp). *Annual review of psychology*, 62:621–647, 2011.

[197] Marta Kutas. In the company of other words: Electrophysiological evidence for single-word and sentence context effects. *Language and cognitive processes*, 8(4):533–572, 1993.

[198] Cyma Van Petten. A comparison of lexical and sentence-level context effects in event-related potentials. *Language and Cognitive Processes*, 8(4):485–531, 1993.

[199] Javier Lopez-Calderon and Steven J Luck. Erplab: an open-source toolbox for the analysis of event-related potentials. *Frontiers in human neuroscience*, 8, 2014.

[200] S. Nishio, H. Ishiguro, and N. Hagita. Geminoid: Teleoperated android of an existing person. *Humanoid robots-new developments. I-Tech*, 2007.

[201] A. Chella, Christian Lebiere, David C Noelle, and Alexei V Samsonovich. On a roadmap to biologically inspired cognitive agents. In *BICA*, pages 453–460, 2011.

[202] R. Sorbello, S.Tramonte, Carmelo CalÃ, Marcello Giardina, Shuichi Nishio, Hiroshi Ishiguro, and A. Chella. An android architecture for bio-inspired honest signalling in human-humanoid interaction. *Biologically Inspired Cognitive Architectures*, pages –, 2017.

[203] Alex Pentland and Tracy Heibeck. *Honest signals: how they shape our world.* MIT press, 2010.

[204] Alex Sandy Pentland. Automatic mapping and modeling of human networks. *Physica A: Statistical Mechanics and its Applications*, 378(1):59–67, 2007.

[205] R. Sorbello, A. Chella, Marcello Giardina, Shuichi Nishio, and Hiroshi Ishiguro. An architecture for telenoid robot as empathic conversational android companion for elderly people. In *Intelligent Autonomous Systems 13*, pages 939–953. Springer, 2016.

[206] A. Chella, R. Sorbello, G. Pilato, G. Vassallo, G. Balistreri, and M. Giardina. An architecture with a mobile phone interface for the interaction of a human with a humanoid robot expressing emotions and personality. *AI* IA 2011: Artificial Intelligence Around Man and Beyond*, pages 117–126, 2011.

[207] V. Gentile, S. Sorce, and A. Gentile. Continuous hand openness detection using a kinect-like device. In *2014 Eighth International Conference on Complex, Intelligent and Software Intensive Systems*, pages 553–557, 2014.

[208] S.M. Anzalone, F. Cinquegrani, R. Sorbello, and A. Chella. An emotional humanoid partner. In *Proceedings of the 1st International Symposium on Linguistic and Cognitive Approaches to Dialog Agents - A Symposium at the AISB 2010 Convention*, pages 1–6, 2010.

[209] R. Sorbello, A. Chella, Carmelo Calí, Marcello Giardina, Shuichi Nishio, and Hiroshi Ishiguro. Telenoid android robot as an embodied perceptual social regulation medium engaging natural human-humanoid interaction. *Robotics and Autonomous Systems*, 62(9):1329 – 1341, 2014. Intelligent Autonomous Systems.

[210] James A Russell. A circumplex model of affect. *Journal of personality and social psychology*, 39(6):1161, 1980.

[211] Mohammad Soleymani, Micheal N. Caro, Erik M. Schmidt, Cheng-Ya Sha, and Yi-Hsuan Yang. 1000 songs for emotional analysis of music. In *Proceedings of the 2Nd ACM International Workshop on Crowdsourcing for Multimedia*, CrowdMM '13, pages 1–6, New York, NY, USA, 2013. ACM.