EasyControl
Ambient Control for Tetraplegics
David Magro Graça

Dissertação para obtenção do grau de Mestre em
Engenharia Informática e de Computadores

Júri
Presidente: Prof. Dr. Paulo Ferreira
Orientador: Prof. Dr. Joaquim Armando Pires Jorge
Vogal: Prof. Dr. João Manuel Brisson Lopes

Novembro 2008
Acknowledgements

Although this work showed some promising results, there were some tough times, especially at the beginning. During these times many persons helped me either with intellectual brainstorming or with simple support. I would like to thank my family for their comprehension and support, for all the times I came home tired, unhappy and without patience. Family and friends, especially Miguel Faneca, were the refuge that allowed me to relax, to talk and unburden all my troubles and worries.

I would like also to thank my girlfriend, Cecília, for her patience when my mind wasn’t where it was supposed to be, for all the care I received, for our brainstorming sessions and specially for being there whenever I needed. I know it wasn’t easy for you too and that I’m a very lucky man for having You by my side. I can’t imagine how it would have been without all your care and affection. You are very special to me.

I know that, although this dissertation wasn’t easy and demanded a lot of work and time from me, it wasn’t easy for them too. They were the ones who supported me when I needed and that made me happy when nothing seemed to be capable of making me smile. A very Special Thanks to All.

Other people helped me and gave me strength to continue. Tiago Guerreiro followed closely my work by not only supervising it but also by helping, understanding and discussing it in a pro-active way. Our discussions expanded my limits and opened my mind in many ways. To him special thanks for the time he spent with me when I seemed confused, for the beginner mistakes that I made and specially for being present whenever I had a doubt, regardless of his time constraints.

To Professor Joaquim Jorge I would like to thank for his ideas when problems appeared and solutions were complicated, for his support on the system implementation and for the proposing such an interesting research topic that became this dissertation. Without him this dissertation wasn’t possible.

Last but not the less important, I would like to thank all the users that participated in this work, especially Pedro Freitas who followed all my work being available for all the tests done. I would like also to thank Nuno Cabruja. All participants showed me strength and willpower motivating me to do an even better work.

Thank You for all the support and comprehension.
Abstract

Technology is becoming more important in our everyday life. Through it our quotidian obstacles can be overcome more easily. Indeed, it's so important that it has become part of our lives, controlling our standard of life and social acceptance. However, some people are deprived of this technology by not being able to use their control mechanisms.

Tetraplegics are among the most limited in terms of motor-skills, having serious difficulties controlling their homes. Taking into account their needs and limitations, a solution was proposed with the aim of developing suitable interaction schemes. The solution used the developed Slash Interaction for all gesture-based input devices (laser and touchpad) and Speech as a substitute for, or to complement, gestures, creating a multimodal interface. The system was intended to be easily adaptable to changes in inputs and in the environment, while using low-cost technology to address the user’s needs. The main objective of this work was to “Return to tetraplegics the ownership of their houses”.

In this document we present an approach that can give back some household control to tetraplegics by enhancing the interaction between the users and the system. In user studies, the interaction schemes showed high user acceptance, performance and accuracy.

Keywords

Assistive Technologies, Tetraplegia, Multimodality, Speech, Gestures, Domotic System.
A tecnologia está a tornar-se cada vez mais importante no nosso quotidiano, permitindo-nos alargar os nossos horizontes. De facto, ao facilitar a nossa vida, tornou-se parte dela, controlando o nosso nível de vida assim como a nossa aceitação social. Contudo para controlar estas tecnologias é necessária alguma destreza motora, o que para algumas pessoas pode representar um problema. Ao não controlar estes mecanismos a capacidade podemos até limitar a nossa qualidade de vida.

Os Tetraplégicos estão entre os mais limitados em termos de capacidade motoras tendo sérias dificuldades em controlar o seu ambiente, que geralmente é a sua casa. Tendo em conta as suas necessidades, uma solução foi proposta que apontou para o desenvolvimento de métodos de interacção adequados. A solução usou o método de interacção desenvolvido, Slash, para todos os aparelhos de captura baseados em gestos (laser e touchpad) e fala para complementar ou substituir os gestos, criando uma interface multimodal.

O sistema foi baseado numa abordagem fácil de adaptar às capacidades dos utilizadores, através dos aparelhos de entrada de dados, e às mudanças no ambiente, usando em ambos os casos tecnologia de baixo custo.

Neste documento é apresentada uma abordagem que pode restaurar alguma da autonomia para este tipo de utilizadores, através de uma interacção mais eficiente entre o sistema e utilizador. Através de estudos de utilizador, o sistema demonstrou ter uma alta aceitação de utilizador assim como eficiência e precisão.

**Palavras-Chave**

Tecnologias Assistivas, Tetraplégia, Multimodalidade, Fala, Gestos, Sistema Domótico.
# Table of Contents

Acknowledgements \hspace{1.5cm} i

Abstract \hspace{1.5cm} iii

Resumo \hspace{1.5cm} v

List of Figures \hspace{1.5cm} xii

List of Tables \hspace{1.5cm} xiv

1 Introduction \hspace{1.5cm} 1
   1.1 Motivation \hspace{1.5cm} 2
   1.2 Home Control for Tetraplegics \hspace{1.5cm} 3
   1.3 Contributions \hspace{1.5cm} 3
   1.4 Publications \hspace{1.5cm} 4
   1.5 Document Outline \hspace{1.5cm} 4

2 Background \hspace{1.5cm} 5
   2.1 Tetraplegia \hspace{1.5cm} 5
      2.1.1 Spinal Cord Injuries and Tetraplegia \hspace{1.5cm} 6
      2.1.2 Aimed Tetraplegia Segment \hspace{1.5cm} 7
   2.2 Assistive Technologies for Tetraplegics \hspace{1.5cm} 8
      2.2.1 Nurses and Caregivers \hspace{1.5cm} 8
      2.2.2 Switches and Pointers \hspace{1.5cm} 9
      2.2.3 Sound/Speech \hspace{1.5cm} 9
      2.2.4 Gaze and Motion Tracking \hspace{1.5cm} 11
      2.2.5 Breath Interfaces \hspace{1.5cm} 13
      2.2.6 Brain-Computer Interfaces \hspace{1.5cm} 14
      2.2.7 Electromyographic Interfaces (EMG) \hspace{1.5cm} 15
      2.2.8 Evaluation Criteria \hspace{1.5cm} 16
   2.3 Environmental Control for Tetraplegics \hspace{1.5cm} 20
      2.3.1 Robots for Environment Control \hspace{1.5cm} 20
      2.3.2 Smart Homes for the Motor-Impaired \hspace{1.5cm} 21
      2.3.3 Evaluation Criteria \hspace{1.5cm} 23
## 2.4 Overall Discussion

### 3 Home Control for Tetraplegics

#### 3.1 Preliminary studies

- **3.1.1 User-Centered Design**
- **3.1.2 User and Task Analysis**
- **3.1.3 User Interaction in Computer Access**
- **3.1.4 Domotics**
- **3.1.5 Current Scenario**
- **3.1.6 Preliminary Requirements**

#### 3.2 Environmental Control for Tetraplegics

- **3.2.1 Objectives**
- **3.2.2 Aimed Scenarios**
- **3.2.3 System Inputs**
- **3.2.4 Proposed Input Methods**

#### 3.3 Available Work Evaluation

- **3.3.1 Hardware Tools**
- **3.3.2 Software Tools**

#### 3.4 Preliminary Tests

- **3.4.1 Motivation**
- **3.4.2 Related Work**
- **3.4.3 Users**
- **3.4.4 Participants**
- **3.4.5 Apparatus**
- **3.4.6 Procedure**
- **3.4.7 Design**
- **3.4.8 Discussion**
- **3.4.9 Post-test analysis**
- **3.4.10 Slash Interaction Mode**

## 4 EasyHouse: Domotic Home Control for Tetraplegics

#### 4.1 System Overview

#### 4.2 Material Used

#### 4.3 Multimodal Interface

- **4.3.1 Speech Module**
- **4.3.2 Gestures Module**
- **4.3.3 Modality Integration**

#### 4.4 Information Structure

- **4.4.1 Physical Layer**
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Motor Map.</td>
<td>6</td>
</tr>
<tr>
<td>2.2</td>
<td>Nurse giving emotional support to a person.</td>
<td>9</td>
</tr>
<tr>
<td>2.3</td>
<td>Switches and Head-Pointers for the motor-disabled.</td>
<td>10</td>
</tr>
<tr>
<td>2.4</td>
<td>Vista Speech initial grid.</td>
<td>11</td>
</tr>
<tr>
<td>2.5</td>
<td>Mihara et al.[37] non-verbal cursor control.</td>
<td>12</td>
</tr>
<tr>
<td>2.6</td>
<td>Eye Tracker and Eye Mouse.</td>
<td>14</td>
</tr>
<tr>
<td>2.7</td>
<td>Fuzzy Mouse being tested.</td>
<td>14</td>
</tr>
<tr>
<td>2.8</td>
<td>128 electrode EEG device.</td>
<td>15</td>
</tr>
<tr>
<td>2.9</td>
<td>EMG device placed on the user's leg.</td>
<td>17</td>
</tr>
<tr>
<td>2.10</td>
<td>ASPICE Graphical User Interface.</td>
<td>22</td>
</tr>
<tr>
<td>2.11</td>
<td>Soft Remote Control.</td>
<td>23</td>
</tr>
<tr>
<td>3.1</td>
<td>(a) The default interface for a print dialog. (b) A user interface for the</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>print dialog automatically generated for a user with impaired dexterity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>based on a model of her actual motor capabilities.</td>
<td></td>
</tr>
<tr>
<td>3.2</td>
<td>(a) Users with motor impairments often have difficulty acquiring area</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>targets (b) In goal crossing, the need for acquiring a confined area and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>clicking is removed; goals must only be passed.</td>
<td></td>
</tr>
<tr>
<td>3.3</td>
<td>Domotic Overview.</td>
<td>31</td>
</tr>
<tr>
<td>3.4</td>
<td>Example of the laser dot recognizing system using Matrix Interaction Method.</td>
<td>37</td>
</tr>
<tr>
<td>3.5</td>
<td>Image from the testing program.</td>
<td>42</td>
</tr>
<tr>
<td>3.6</td>
<td>Existing movement classification using in every Matrix-based interaction</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>mode.</td>
<td></td>
</tr>
<tr>
<td>4.1</td>
<td>EasyHouse architecture.</td>
<td>46</td>
</tr>
<tr>
<td>4.2</td>
<td>Speech Commands.</td>
<td>49</td>
</tr>
<tr>
<td>4.3</td>
<td>Gesture capture algorithm.</td>
<td>52</td>
</tr>
<tr>
<td>4.4</td>
<td>Filters applied to an image from the webcam.</td>
<td>53</td>
</tr>
<tr>
<td>4.5</td>
<td>Overview of the system layers.</td>
<td>54</td>
</tr>
<tr>
<td>4.6</td>
<td>Transmission of data through electricity using X10 protocol.</td>
<td>55</td>
</tr>
<tr>
<td>4.7</td>
<td>Structure of the most important objects in the system.</td>
<td>56</td>
</tr>
<tr>
<td>4.8</td>
<td>Overview of the Application Layer.</td>
<td>57</td>
</tr>
<tr>
<td>4.9</td>
<td>Overview of the Graphical Interface.</td>
<td>57</td>
</tr>
<tr>
<td>4.10</td>
<td>House and Device XML configuration.</td>
<td>60</td>
</tr>
<tr>
<td>4.11</td>
<td>Configuration panel in the interface.</td>
<td>60</td>
</tr>
</tbody>
</table>
4.12 In this example the red arrows express a directions that has no effects due to the unexistence of objects in the correspondent cells. The green arrows correspond to directions that are acceptable. 61

5.1 Test Environment. ................................................................. 66
5.2 Errors Comparison. ............................................................... 69
5.3 Comparison between average times and the modalities modalities used. ........................................ 69
5.4 Errors per modality. ............................................................... 72
5.5 Errors per Task. ................................................................. 73
5.6 Times taken per modality. ........................................................ 74

A1.1 Endpoints and path measures relative to Task Axis. ................................................................. 88
A1.2 Path variations per device. ........................................................ 89
A1.3 Mean Selection Time, Target Acquisition Time of the devices. .............................................. 90
A1.4 MT a and b determined experimentally per device. ................................................................. 90
A1.5 Laser Times compared statistically. ................................................................. 90

A2.1 Example of interaction. The user will turn on the lamp in the bedroom. The green arrows were here used for better understanding of the movements done to activate an option. .......................... 94
List of Tables

2.1 Comparative Human Computer Interfaces Table. ............................... 19

5.1 Users technological profile table. .................................................. 67
5.2 Time Comparison between (Multi-)Modalities. ................................. 75
5.3 Error Rate Comparison between the Preliminary Evaluation and the Final Evaluation. ................................. 75
5.4 Time Comparison between the Preliminary Evaluation and the Final Evaluation. ................................. 76
Introduction

The World Health Organization (WHO) defines a person with disability (or activity limitation) as a person of any age who is unable to perform independently and without aid, basic human activities or tasks, due to a health condition or physical/cognitive/psychological impairment of a permanent or temporary nature [9]. According to the World Bank estimate about 10% of the world's population is suffering from some kind of disability. With these statistics we can change preconceived ideas about the disabled people because it proves that they're not an occasional exception, in fact they are a big piece of the global population that has serious needs which are frequently neglected and for whom little improvements can have significant results. This percentage will rise since globally the average age is rising, and it's believed that 58.5% of people over 65 have disabilities, with motor disabilities making up a large part of this.

To the motor disabled person, actions like going to the bathroom can require a significant increase of effort without a caregiver's or system's help and that reflects on the person's psychological and physical health. Although the physical damage is, in some cases, severe, the psychological is frequently more severe. Technology nowadays is sufficiently advanced to provide user solutions that tackle some of their needs or facilitate the access to others. In a house there are many appliances, most of them with different strategies for controlling the device, and almost none suited for the physically challenged. This situation imposes a dilemma for people that have a house and live in it, which is the lack of autonomous control over almost every appliance, transforming their house into a caregiver-operated house and not their own house. Nowadays technology can provide us with a number of remote control systems for home appliances, frequently not suited for the physically challenged, which are complicated and few adapt to house modifications.

In this dissertation I will focus on a subset of the physically challenged which are the tetraplegic users because of the serious needs that they have and the lack of existing usable solutions. A tetraplegic is a person that has severe motor constraints in the four limbs usually caused by an spinal cord injury, which is a long, thin, tubular bundle of nerves which is an extension of the central nervous system from the brain and is enclosed in and protected by the bony vertebral column. In this tetraplegic subset limb control or trunk control differs mostly by the vertebrae and the type of the injury. The injury level affects users capabilities, injuries may vary from a very high tetraplegia (vertebrae C1-C2), where the person doesn’t have control over
the diaphragm muscle and therefore can’t breath without help, to a low tetraplegia (C5-C8), where one may sense or have some kind of control over his/her shoulders, arms and wrists. Cognitive capacities remain, although intact meaning these individuals are capable of performing actions if the proper interfaces are used.

This work tries to capture the most relevant systems that aim for the rehabilitation of these people, in the sense that a person using a combination of simple, low-cost systems can re-gain some of his/her old abilities in their homes, restoring some of the tetraplegics autonomy while expanding their action range.

1.1. Motivation

For people without movement limitations this can be a strange subject to work on because they have never experienced the workload of getting from point A to B, typing an SMS on a mobile phone or having a glass of water, when your own body doesn’t respond to your intentions. These tasks, for people with motor disabilities, can represent, by our standards, the same as a day of work or maybe more. So, can a decrease in the complexity of a frequent task be relevant to their day-to-day life? Our environments are constructed and developed without attention to the needs of people with movement constraints, even though, with the life expectancy rate of the developed countries, they make up a great slice of the world population; 58.5 per cent of people over 65 have disabilities, suffering the same problems that constrain their movements [36]. There are some approaches that are aimed at people with high tetraplegia and a few are usable for low tetraplegia which represents most of the population with spinal cord injuries. These people have some control in their limbs but not enough to use existing solutions in a satisfying way. Methods have to be discovered to take advantage of their residual inaccurate movements.

In the case of the tetraplegics, also known as quadriplegics, the movement limitations are severe although they are in perfect cognitive health. For them every basic action must have facilitated access in order for them to have control over it, and when that doesn’t happen a caregiver must provide support to it. This situation makes the caregiver an essential piece of the communication and control such an individual has over their environment and when that piece fails the individual’s whole life fails too. Caregivers have the hard task of doing all the tasks that a normal person does both for the tetraplegic and for themselves, so they therefore have to act for two people.

This close relationship between the tetraplegic and the caregiver might at first seem to be unnecessary for simple, common tasks such as turning on the TV, but tasks like these, due to the frequency that they need to be carried out, represent a major workload to the caregiver while the tetraplegic’s control over their environment vanishes. By enhancing the user’s control over their environment through the use of technology and by taking advantage of vocal and gestural capabilities a solution can be sought that matches those needs, reducing caregiver intervention and improving a tetraplegic’s control.

But can we adapt the environment to meet the user’s needs? The simple answer is ‘Yes’. By providing a common interface to the most frequent devices and integrating them in a system it is possible to achieve a similar quality level as that of a person without any kind of motor disability with the advantage that these systems can also be used by those people.
1.2. Home Control for Tetraplegics

The approach taken makes use of the user’s body movements, taking into account that small movements are extremely inaccurate and directionally inconstant, and that wider movements reveal a more accurate direction and are more stable. This document presents an approach that offers to the tetraplegic a way or ways to control their surrounding appliances without any caregiver help, enabling them to set up their home environment as they please. The system can operate by activating or deactivating electrical appliances and by increasing or decreasing electrical devices like the blinds and the lights. To achieve this goal, a set of modalities had to be studied and systems had to be developed to enhance the communication quality and bandwidth between the tetraplegic and the system. User studies were carried out to characterise needs, physical capabilities and assistive technologies suitable for those needs have been studied.

For the development of the system, context was specified which took into account the user’s accommodations, needs, habits and available technology, and potential system expansion to other types of physically-challenged people’s scenarios provided additional information. This challenge was complex due to differences in the capabilities of each individual, which varied within the tetraplegic subset of the population.

The approach taken focused on a set of input devices that can be used in different situations, some when the user is in bed, others when the user is in a wheelchair. Since every person in the target population can speak without any constraints, speech, with its high dimensionality and high command rate, was a natural choice when compared with other inputs. The voice has some inherent problems due to noise or user acceptance issues. User tests in movement analysis showed some consistency in direction and low accuracy. This led to a stroke-based interaction, the Slash Interaction, which makes use of those aspects, classifying the movements into 8 movement classes (horizontal, vertical and diagonal movements). This method was originally developed for a Laser Pointer interaction due to adaptability factors but was rapidly expanded to all gesture-based input devices, laser and touchpad.

The control of appliances over the electrical wiring of the house used the X10 protocol being a implementation developed due to the lack of existing usable solutions on the market. The X10 uses the house power line to create a domotic network. It is a low-cost and easy adaptable technology that can be adjusted to the surrounding environment.

The overall system had speech, gestures (through laser and touchpad) and appliance control over the electrical wiring, allowing the user to chose how to interact with the system. With this system, a tetraplegic user proved that he/she can remotely control appliances successfully, relieving much of the caregiver’s work while increasing user satisfaction and autonomy.

1.3. Contributions

The main contribution of the research is the design and development of a system that allows tetraplegic users to remotely control house appliances naturally and accurately. To achieve this, various steps had to be taken and each one also represented a contribution. Below there are the contributions that this work accomplished:
• *EasyHouse: A Multimodal Domotic System for the Tetraplegic*, having a user-centred approach that offers the possibility to tetraplegics to control remotely the electrical devices in their house. This multimodal prototype uses Speech and/or Gestures, through any input device that moves the Windows cursor, and the X10 domotic protocol to control the appliances. It was evaluated in many stages of development taking into account the user's needs and showed promising results either in accuracy or in performance.

• *Slash interaction*, was a new concept created for gesture recognition, input device independent, that was based on the tetraplegics arm/hand movement characteristics. It classifies fast, undirected movements into 8 movement groups which correspond to horizontal, vertical and diagonal movements.

• A study on Pointing Tasks and Home Network Control Systems for the Motor-Disabled; this analyzed the advantages and disadvantages of the existing approaches, pointing out the problems and results of each approach and analyzing their suitability to the scenarios and user environments.

• A library that can operate remotely any electrical devices; based on the Home Network studies an approach was selected and due to the non-existence of suitable implementations a library was constructed to provide a computer interface to the electrical wiring of the house.

1.4. Publications

The work presented in this dissertation resulted in one original publication accepted in peer-reviewed Scientific Conferences. Below I present it:


1.5. Document Outline

To capture technologies suitable for the tetraplegic a study was made that surveyed the existing assistive technologies and the home network technologies available, Chapter 2. To develop a system that satisfies the needs of the tetraplegic, studies had to be carried out that better characterised the target population, providing clear objectives and requirements that became the pillars of the dissertation while proposing an approach that tackled the limitations found, Chapter 3. Based on the analysis done in Chapter 3, the ideas were developed and the system was built, Chapter 4. The developed prototypes were then tested with users to assess and/or validate the choices followed. The assessment was based on the objectives and requirements created in Chapter 3 using specific metrics that allowed the comparison between testing stages, Chapter 5. The last chapter present the conclusions taken from the work done and proposes a set of future developments in the system.
This chapter presents some of the existing work which was analyzed and discussed, creating a path that directed all the approach in this dissertation. It gives an outline over the existing technologies and solutions available but also over the tetraplegic users because technology can’t work without being fitted to its users and all discussions and assessments must be based on that principle. Making assessments without having the users profile into account could only lead to flaws in the discussion and possible wrong assumptions throughout the dissertation. Therefore this chapter covers both factors, technology and user profiles, creating contextual discussions which naturally will lead to a path.

2.1. Tetraplegia

Tetraplegia, also known as Quadriplegia, is a symptom in which a human experiences paralysis affecting all four limbs, although not necessarily total paralysis or loss of function. It is caused by damage to the brain or the spinal cord at a high level, in particular spinal cord injuries secondary to an injury to the cervical spine. The injury, known as a lesion, causes victims to lose partial or total mobility of all four limbs, meaning the arms and the legs. Typical causes of this damage are trauma (such as an auto accident, gunshot wound, fall, or sports injury) or disease (such as transverse myelitis, polio, or spina bifida). It is possible to suffer a broken neck without becoming quadriplegic, such as when the vertebrae are fractured or dislocated but the spinal cord is not damaged. Conversely, it is possible to injure the spinal cord without breaking the spine, such as when a ruptured disc or bony spur on the vertebra protrudes into the spinal column.

Vertebral segments

There are 7 cervical (neck), 12 thoracic (chest), 5 lumbar (back), and 5 sacral (tail) vertebrae. The thoracic vertebrae are defined by the spinal cord segments which are not necessarily situated at the same vertebral levels. For example, while the C1 cord is located at the C1 vertebra, the C8 cord is situated at the C7 vertebra. While the T1 cord is situated at the T1 vertebra, the T12 cord is situated at the T8 vertebra. The lumbar cord is situated between T9 and T11 vertebrae. The sacral cord is situated between the T12 to L2 vertebrae, Fig. 2.1.1.
2.1.1. Spinal Cord Injuries and Tetraplegia

In the US 92% of the spinal cord injuries result in Tetraplegia being the motor vehicle accidents the most frequent cause of spinal cord injuries (44%) which represents 7800 spinal cord injuries per year only in the US, according to the National Spinal Cord Injury Association (NSCIA)\(^1\). A Spinal cord injury causes myelopathy or damage to white matter or myelinated fiber tracts that carry sensation and motor signals to and from the brain. It also damages gray matter in the central part of the spinal, causing segmental losses of interneurons and motorneuron. Since the spinal cord is the part of the body that is responsible for the control of the body, injuries in it cause immediate loss of sensitivity and/or motor control in some parts of the body. The cells that compose the spinal cord are neurons which are only defined in our early ages having residual cell replication in latter stages, making the injuries permanent.

![Motor Map](image)

(a) Relation between injuries and body parts.  
(b) Relation between injuries and muscle control.

Figure 2.1: Motor Map.

Cervical (neck) injuries usually result in full or partial Tetraplegia, depending on the location of the injury. Persons with a spinal cord injury at the cervical level may retain some amount of function or sensibility but are otherwise completely paralyzed. Tetraplegics have injuries from C1 to C8. Figure 2.1(a) and 2.1(b) express the relation between injuries and body part control. Bellow there is the relationship between vertebrae injury and motor control in Tetraplegics:

- Cervical (neck) injuries usually result in quadriplegia;
- C1 to C4 injuries may require a ventilator for the person to breathe;
- C5 injuries often result in shoulder and bicep control, but no control at the wrists or hands;

• C6 injuries generally yield wrist control, but no hand function;
• C7 and T1 injuries can straighten their arms but still may have dexterity problems with the hand and fingers;
• Thoracic level injuries and below result in paraplegia, with the hands not affected;
• T2 to T8 injuries have control of their hands, but poor trunk control as the result of lack of abdominal muscle control;
• T9 to T12 injuries allow good trunk control and good abdominal muscle control. Sitting balance is very good;
• Lumbar and Sacral injuries yield decreasing control of the hip flexors and legs.

Besides a loss of sensation or motor functioning, individuals with SCI also experience other changes. Some of the symptoms that occur are: Incontinence of the bowel and bladder, Erectile dysfunction, Fertility may be affected in SCI men, Wide and rapid fluctuations in body temperature, Inability to sweat below the level of injury, Low blood pressure and Chronic pain.

Very high injuries (C1-C2) can result in a loss of many involuntary functions. These injuries often affect the ability to breathe (requiring breathing aids such as mechanical ventilators or diaphragmatic pacemakers) and the ability to regulate blood pressure effectively.

Within the Tetraplegics there are 3 distinct categories distinguished by level of injury: Very High Tetraplegia (C1-C2), High Tetraplegia (C2-C5) and Low Tetraplegia (C5-C8).

Beside the characterization by injury level, there is also another characterization based on the motor and sensory control over both parts of the body. Those categories are: Complete or Incomplete injuries. The Complete injuries are characterized by the loss of motor control and sensory function bellow the place of the injury, both sides equally. Incomplete injuries there may be sensory function in some part bellow the injury or motor complete/partial control in one of the limbs bellow.

2.1.2. Aimed Tetraplegia Segment

According to NSCIA, 92% of the SCI result in Tetraplegia being the majority of those injuries classified as High Tetraplegia, characterized by injuries between C2 and C5 vertebrae. To understand the difficulties inherent to this type of Tetraplegia in terms of technology, we must discuss some of the problems in Very High and Low Tetraplegia.

Very High Tetraplegia is a characterized by severe health problems due to the inability to move and control almost every part of the body including the diaphragm muscle, which control our breath. Movement if possible can only happen in the head region, increasing the difficulty of any gesture-based control interface. A person in this condition can have preserved mental capacities but can’t interact with most of the surrounding devices, creating a technological barrier, difficult to surpass. The limitations with these tetraplegics such as the inability to control the breath and/or mental problems impose other types of restrictions when regarding device interaction.

Low Tetraplegia is characterized by a relative high degree of motor control in the arms and hands, although with reduced dexterity. The degree of control and accuracy of these subjects is enough to control,
some difficulty, the most frequent gesture-based control interfaces, easing the interaction with the surrounding
environment and enhancing their communication skills, because they are able to use regular devices. The
importance of easier input modes has therefore a lower importance since they already have enough dexterity
and muscular control to interact with the regular devices.

In the case of High Tetraplegia a dilemma exists due to the preserved mental skills and the insufficient level
of motor control to interact with regular gesture-based control interfaces. These persons are therefore in a
zone which the technology used by able-bodied persons can’t be applied although the mental skills are similar.
Therefore the aim of these studies is to discuss, find and fulfil this gap enabling the High Tetraplegics to once
again interact with their home appliances. Creating ways to simplify these interactions, the gap created by the
lack of motor-skills can be reduced which in the case of High Tetraplegics is enough to provide and increase
allowing them to use regular household appliances.

2.2. Assistive Technologies for Tetraplegics

Assistive technology is a generic term that includes assistive, adaptive, and rehabilitative devices and the
process used in selecting, locating, and using them. AT promotes greater independence for people with dis-
abilities by enabling them to perform tasks that they were formerly unable to accomplish, or had great difficulty
accomplishing, by providing enhancements to or changed methods of interacting with the technology needed
to accomplish such tasks. Although, Cook and Hussey [12] report this term is usually not used for rehabilitative
devices and for devices that non-disabled people find useful. According to disability advocates, technology is
often created without regard to people with disabilities, creating unnecessary barriers to hundreds of millions
of people.

Every system needs to interact with its users to achieve its goals however this doesn’t mean that the pro-
cess is always done by choosing the right interface to achieve maximum effectiveness because that will lead
us to solutions that will only fit specific and controlled situations. Having into account various real situations,
such as when the user is lying down, in the wheelchair in house or outside, the technologies must be eval-
uated regarding uniform criteria to discover one or more solutions capable of being viable and useful to its
users. Bellow I briefly explain some of the existing technologies specifying its problems and advantages.

2.2.1. Nurses and Caregivers

Nurses and Caregivers represent the main support system that helps and maintains the life quality of the
elders and the disabled by providing specific services and assisting them in their daily life which, in the case
of persons with severe motor or cognitive impairment, represent an intense and demanding task. Usually,
when users are capable of living at home, they have medical knowledge although not acquired by school,
depending on the user to tell them what to do. In severe motor-disabled cases it becomes impossible to
satisfy the tetraplegics user hygiene needs at home, therefore they stay in a hospital where they are taken
care by specialized personnel with higher qualifications.
There are many areas that these caregivers work on, from emotional support, hygiene tasks, transportation to physiotherapy, health monitoring and treatment these multitask persons do it all. This multitasking is expensive in a economic or emotional view because of its intensive workload on the disabled persons.

2.2.2. Switches and Pointers

The most common devices used to control a computer interface are the switches. These are electronic devices that send a binary output, i.e. On/Off output, which represents a small amount of information, Figure 2.3(b). The movement required to activate such devices depends of the type of the switch but usually only a small amount of movement is needed. There is a large number of switch types differing in movements to be operated, they can be operated by hand, fingers, chin, head, tongue, etc. This adaptability made these devices one of the most common nowadays, not only for the physically challenged but also for the able-bodied, due to their simplicity and robustness (Keyboard and the Mouse per example). Further to their wide usage, there are some enhancing techniques that aimed to the motor-disabled. The scanning interfaces are a concept used to facilitate the selection of an option, these interfaces usually highlight a option at a time which is selected when the user presses a button.

Other approaches take advantage of the head motion to activate buttons, Figure 2.3(a). It’s a attached device that extends the user range of action by using a stick. This device can be adapted to use the chin, cheek or teeth. These variations are used in many scenarios, for instance, in the wheelchair the guidance joystick can be replaced by chin joystick with performance improvements for the motor-disabled.

2.2.3. Sound/Speech

Sound and Speech are the main forms of communication in the world, they are characterized by a high portability since they can be reproduced remotely without change of meaning and by a universal availabil-
ity. Considering the motor-disabled the Speech/Sound modalities maybe the only remaining intact modality available although that depends of the degree of the injury. It's naturalness and high dimensionality, input rate and easy conjugation with other modalities make these modalities some of the most promising for the target population.

There have been many approaches that focused on the adaptation or enhancement of these modalities to fit the motor-disabled population needs. To access information in a computer a person must have a high degree of motor control to operate the input devices, mainly the keyboard and the mouse, which clearly is difficult to achieve due to physical limitations. The use of speech to emulate the directional fine-grained movements was, since the beginning, found as inadequate and tended to be slow and error prone. Manaris and Harkreader [35] developed a system called SUITEKeys which was a continuous speech engine that made the direct mapping between user speech and keyboard/mouse operations, these researchers concluded that “speech isn’t the best modality for all human-computer interaction tasks” when compared with the keyboard and mouse but “allows universal access to computing devices”. Mauro, Gori, Maggini, and Martinelli [14] discussed the design of a voice-controlled mouse, navigation was supported using a set of commands that required users to learn non-intuitive mappings between commands (e.g., “Move left”) and utterances that caused commands to be executed (e.g., “A”) no data were provided regarding the efficacy of this solution. Karimullah and Sears [27] confirmed that basic direction-based solutions that result in fixed-speed continuous cursor movements could allow users to accurately select targets that are sufficiently large. However, the time required to complete these tasks increased significantly as the distance to the target increased, error rates increased dramatically as the size of the target decreased, and user satisfaction ratings were not encouraging. Diagonal movements highlight yet another limitation of this solution.

Dai et al. [13] adopted a grid-based approach that recursively splits the speech selected area into a 3x3 matrix until the user issues the selection order. The systems works by numbering every cell in the matrix with a number (1-9) which the user can select by speech to perform a zoom that splits the selected cell also in a 3x3 matrix, this process is done until the user issues the select action, Figure 2.4. Tests showed a 33% de-
crease when compared with results of other studies with large targets and up to 55% increase in time savings for small targets. This approach is similar to the encountered in Windows Vista.

![Figure 2.4: Vista Speech initial grid](image)

Igarashi and Hughes [24] used a different approach, instead of recognizing words and phrases they used the low-level characteristics of voice/sound such as pitch, volume and the duration of sound. They found that the continuously vocal sound tires the throat and is a unnatural way of interaction. The goal of these approaches isn’t the replacement of the traditional method, it’s instead a complement, they stated that “Voice-as-sound techniques complement traditional speech recognition interfaces rather than replacing them, by allowing the user to directly adjust system parameters.”. Mihara et al. [37] developed a multiple cursor system also based in the matrix selection method, selecting the columns and then the rows by saying the correspondent cursor number for large grained areas and then non-verbal vocalizations, such as “Ahhhhh”, to make finer-grained adjustments. The choice of the use of non-verbal vocalizations was due the fact that speech commands carry a delay when used to perform selections or fine-grained adjustments. The system uses then discrete signals to locate the target and after continuous signals that move the cursor slowly making adjustments, Figure 2.2.3.

### 2.2.4. Gaze and Motion Tracking

There have been several approaches that captured the user focus or intention to ease a task. Since the eyes are the organ that we use to focus anything, if their movements are analyzed we can assess where the user is focusing. Eyes and hands/gestures tend to work together to achieve a target in WIMP interfaces but only the gestures are used to send information, this is the main reason that lead the researchers to try to overcome this limitation. To study the eye-movements many of the approaches used Electrooculography.

**Electrooculography** (EOG) is a technique for measuring the resting potential of the retina, creating a graphic called electrooculogram, by analyzing the surrounding muscles. This technique uses pairs of electrodes placed either above and below the eye or to the left and right of the eye. If the eye is moved from the center position towards one electrode, this electrode detects the positive side of the retina and the opposite
electrode detects the negative side of the retina. Consequently, a potential difference occurs between the electrodes. Assuming that the resting potential is constant, the recorded potential is a measure for the eye position.

Shumin Zhai [65] studied the use of eye movements to reduce context and with it improve reliability and performance. Several of his studies are compared in spatial tasks, the MAGIC [66] studies showed that the "eye as a primary perceptual organ doesn’t function very well" due to difficult fine-grained movements although it was accurate and easy at other movements. Therefore to achieve a target, it used eye gaze to select a wider area and gestures for adjustments, the selection was made by dwell time or blink. Another approach was EASE [60] that used eye gaze to select the closest target possible, the selection being made by a button. This work reduced the cognitive load and the distraction of the user.

Kwon and Kim [31] used EOG signals generated by eye movements and blinks for emulating the PC mouse. For the user's convenience, electrodes were positioned on five specific points on the glasses' frame, in contact with skin and requiring no electrolyte gel. The estimated information is sent to a PC through a wireless RF linkage emulating a wireless mouse. With this they could control basic Windows functions and even play TETRIS using only directional commands (up, down, left, right). Bien et al. [6] designed a system that used CCD cameras to detect the iris, and found that the recognizing system could perceive directional intention and translate that into system actions. The researchers then used that information to direct and control an automated wheelchair.

Other Gaze tracking devices use simpler approaches that reduce aesthetics problems and portability by, instead of looking directly at the eye, looking at head direction and motion. These technologies appeared due to the fact that generally, a person looking directly at an intended object aims the head and the eyes at the same or similar point. Head Optical Pointer is one technology that makes uses of such an assumption.
**Head Optical Pointers** are devices that emit light, generally infrared, which is analyzed and used to issue commands. It's similar to the Head Pointers with the difference that the working range is much higher and there is no need for physical contact to make selections. The light emitted is captured by a set of sensors, direction is discovered by mathematical models and the user head position is assessed.

Chen et al. [10] developed an eyeglass-mounted infrared system to control home appliances for tetraplegic users. The users could select the desired appliance from a board by aiming at the target with the head and staying there a predetermined amount of time. Their studies showed great accuracy, around 85.9%, and an average time cost of 57 secs for the able-bodied and 66 secs for tetraplegics.

**Head Tracking** is another concept that generally uses images to detect human contours or colored labels. The complexity of these systems is focused on the image processing algorithms which have the responsibility of sending correct information about the user's head features such as position and direction.

Head tracking devices are devices for people with preserved neck muscle strength and are based on a set of sensors that monitor a set of characteristics, providing a natural interface for the user. This technique is widely used [64] in spinal cord injury cases because of its low implementation costs and easy adaptability to existing systems such as mouse or other directional systems [19]. These tracking devices need to be attached to the user, and therefore have lower usability and ergonomy, the larger they are the greater being the discomfort level of the user.

To increase aesthetics and usability some approaches used cameras to track marked signs in the user's head avoiding a certain level of discomfort but still not making these techniques very usable.

Bates and Instance [3] developed an interface using a magnification technique directed by the eye and showed that eye-based performance was enhanced, which, combined with favourable user satisfaction tests, proved that it could be a viable way to make the eye mouse systems usable. In Bates Phd Thesis [5] he managed to achieve results that approached those for the mouse using a support modality for object manipulation, buttons, and also object size magnification, Figure 2.6(a). Tests showed that eye and head tracking have some disadvantages due to limitations in the recognition systems, Bates and Instance [4] did a comparison study between those devices and results showed that pointing with these systems is error prone due to involuntary movements by the subjects, the eye being the more inaccurate, although it is also less tiring and faster.

### 2.2.5. Breath Interfaces

Sip and Puff interfaces are the most common among very high tetraplegic users. It is primarily used by people who do not have the use of their hands, and usually this interface is for wheelchair control. The mouth-controlled input provides a simple and effective way to control mouse movement. Movement and operation of this joystick is similar to that of a "Mouthstick". Mouse button clicking is accomplished with the help of the sip or puff function of the joystick.

Other approaches followed using "Sip and Puff" input by capturing air flow presence and air velocity.
(a) The eye mouse in operation. (b) Example of eye racking using infrared sensors.

Figure 2.6: Eye Tracker and Eye Mouse.

Surdilovic [53] captured these signals and transformed them into Morse Code which allowed the user to select the desired actions, Figure 2.7.

Figure 2.7: Fuzzy Mouse being tested.

Breath alone can also be used as an input, Shorrock et al. [50] developed a prototype using a optical mouse that monitors the changes in the diaphragm when the person breathes. The system used a one-dimensional text entry program for testing and achieved a better performance when compared with a Morse Code input method. Despite the good results the system had, it had limited application to domains other than text-entry.

2.2.6. Brain–Computer Interfaces

The most accepted definition of BCI is "Brain–Computer interfaces give their users communication and control channels that do not depend on the brain’s normal output channels of peripheral nerves and muscles" [63]. There are many types of BCI, but due to several factors, Electroencephalography (EEG), similar to EOG, is the only one explained here because it is the most non-invasive interface. Other advantages of this interface
are its fine temporal resolution, ease of use, portability and low set-up cost. Vaughan et al. [58] did a study in which he explains the advantages and disadvantages of each approach. EEG is a technology susceptible to muscle noise, a feature that creates a substantial barrier that can only be tackled with extensive training. It captures the Central Nervous System signals and transforms them into the intended product activity. There are other problems, which will hopefully be overcome with technology, that make this approach only viable for severely disabled people, such as the low rate of signal recognition and the high-level commands that can be extracted. It requires the use of a set of electrodes in the user’s head to capture those signals, Figure 2.8, reduced mobility and aesthetics.

Although it’s still a limited technology many projects have already used it for distinct usages. Moore et al [39] used BCI for controlling hand movements, Pfurtscheller et al. [46] navigated in a house, Milan et al. [38] used a robot for house navigation like the AIBO robot used in the ASPICE project [11].

### 2.2.7. Electromyographic Interfaces (EMG)

This technique detects the electrical potential generated by muscle cells. It detects when the muscle cells contract and translates those signals into intended actions of the user, Figure 2.9. A user that has limited limb movement but still controls his arm muscles can express himself in the same way as he/she did before the injury. This technique can be more useful than other existing techniques such as voice, pointers, keypad, etc. because it requires less movement than using one’s shoulder or neck. Working with these interfaces will only be possible if a learning process is undergone to identify and boost recognition rates.

Nagata et al [41] developed a system that used the three electromyography (EMG) signals of the arm instead of the mouse operation by the hand. The signals which are generated by the movement of the wrist flexion, wrist extension, pronation, supination and grasp are then used to detect live movement of the mouse. In order to distinguish five movements (left, right, up, down, selection) they placed a set of electrodes in the users’ arms. Their system needed a great amount of training to make the correct movement classifications and was aimed at use by amputees.

Jeong et al. [26] developed an EMG-based mouse control method for helping tetraplegics to use the computer by teeth-clenching. The device was attached to the user’s forehead and the system worked by at first choosing the desired direction to move the pointer using left or right teeth-clenching. Then the user

![Figure 2.8: 128 electrode EEG device.](image)
moves the pointer toward the determined direction by clenching all teeth once for more than one second. For stopping movement of the pointer, the user needs to clench all teeth once for less than one second. This approach can provide a great deal of convenience and can adapt to a wider user population. Results showed that the system can be used with tetraplegic users although it was many times slower than a conventional mouse.

2.2.8. Evaluation Criteria

Each technology has particular characteristics which are aimed at a specific user group. To make any kind of comparison a common set of evaluation topics had to be created, giving a clear background over future choices in assistive technologies. The evaluation topics were chosen by their capacity to segment the target user population in an intelligible and objective way. The topics chosen were:

**Potential user range.** Every assistive technology is aimed at a user group and its specific needs, limitations and capacities. This relation user/modality sometimes makes it impossible for other users to use the same approach. Taking into account each user’s limitations, severe in the tetraplegic’s case, not every assistive technology can provide the same kind of results for every type of user, thereby reducing its application. By characterising the requirements needed to operate a certain modality/assistive technology and mapping them with each user group's motor and sensory abilities an assessment can be made that points out the potential user groups that could use the technology.

**Dimensionality and Input Speed.** These are topics that try to characterise the amount/rate of information that can be delivered. They specify the vocabulary (Dimensionality), which is the number of different commands or types of data that can be gathered, and the speed that the data can be provided (Input Speed). If the information is gathered quickly and with various types of data, there is a high possibility that meaningful/data supported actions can be made.

**Accuracy, Robustness and Repeatability.** This set of topics try to characterise the reliability of the approach by explaining the situations that could provoke failure or error. With this set at least a part of the consistency and maturity state of the technology can be assessed.

**Ease of use.** All technologies/modalities are used or tried by users, although their application may vary. To use a technology a set of arrangements have to be made, either physical preparations such as connecting the system components or training users on the system to achieve the expected performance. All these preparations affect system utilization. In all systems there are required actions, independent of the modalities used, which have a difficulty level that depends on the user. By evaluating the frequency of user interaction and the level of difficulty and aggregating that with the preparations needed, one can get an idea of how easy the technology is for users.

**Aesthetics, Hygiene and Acceptance.** For a technology to be applied to a user, some preparations have to be made either with the technology or the devices used. Some modalities require the devices used to be attached to the user's body which may constrain movements, have low aesthetics or imply hygiene issues. These constraints may make it difficult or awkward to use the modality in public places or be intrusive in a way
that increases the user limitations, discouraging user acceptance.

**Mobility Adequacy.** Many technologies are made to work in a specific environment and when the environment changes problems arise and results decrease. This topic characterises the ability of the technology to adapt to environments other than the environment which it was conceived for.

**Availability and Cost.** Technologies may require specific types of hardware/software to operate which have to be acquired. By characterising the cost inherent to the use of the technology and specifying the availability and general usage of the technology we can assess the economical barriers.

**Discussion**

To discuss and compare input devices or modalities a set of common criteria must be defined, establishing objective metrics that can be used for comparison between technologies. This normalisation is advantageous for this work and for future works because it allows other technologies to be assessed. The common ground established allows and eases the whole analysis process, directing the choice of the best technology taking into account the user context.

**Potential users range.** Switches and Pointers don’t require a specific body part to be operated. They can be operated by nearly any body part using anything from residual to wide movements. This potential can be easily adapted to the user’s abilities, facilitating the widening of the user range, if necessary. As well as Switches, EMG, Motion Tracking, Speech/Sound and Breath interfaces also rely on muscular movements that, although small or residual, can be found in most motor-disabled users. Only the top-level spinal cord injuries can invalidate such interfaces due to the vegetative state imposed by them, requiring in most cases even assistive breathing to sustain the user’s life. When we are talking about user range, particularly with motor-disabled users, we have to take into account severe movement constraints, so movement-based interfaces can be presented as error prone and reduce the user range. Some of these interfaces tackle this problem
by using alternate parts of the body such as the brain (BCI). Since the human brain can be controlled with some training and therefore present some voluntary activity it can be adapted from High Tetraplegia users to Able-Bodied users making this the interface with the widest user range.

**Accuracy, Robustness and Repeatability.** Taking into account just the interfaces, Switches or Head Pointers can be presented as the most accurate since they use buttons, which have only two states, On or Off, so the user’s intentions are directly mapped to the interfaces. When considering the target user group, which is characterised by lack of motor control and coordination, the head, shoulder and tongue motions may be the most accurate due to better motor control over those body parts in most users with severe injuries. Sound/Speech interfaces nowadays are relatively accurate, though their performance decreases when in noisy environments such as public places. Sometimes they require a small training session to boost the accuracy and robustness of the interfaces. On the other hand, their recognition rate is high. Gaze/Motion tracking interfaces like eye/head tracking are relatively accurate although these systems require a calibration period and over time the performance will decrease [34], making their robustness strong over a short time, but over longer periods weak or unstable.

**Dimensionality and Input Speed.** The amount of data, disregarding its quality, can be given by the multiplication of Dimensionality and Input Speed, leaving to the researcher the job of parsing and extraction of meaningful information. Speech/Sound is the richest interface in terms of vocabulary (the amount of different sounds that can be extracted and have important information), the input speed that it is generally delivered is medium, similar to the speech rate of a normal person, although some approaches used non-verbal vocalisations which reduce the amount of information but increase data quality. Gaze and Motion tracking interfaces outweigh the reduced dimensionality with the speed rate which is very high, creating a large amount of information, although much of it requiring complex data processing/classification stages. EMG such as Switch interfaces depend on the number of electrodes/buttons used, and the users’ ability to perform movements, which is a limitation in the motor-disabled.

**Ease of use.** Switch interfaces use simple movements to activate a set of buttons, and the devices required can be adapted to the parts of the body which can perform better, easing the process for the user and increasing accuracy. Due to the physical constraints of users a movement-based interfaces may have limited functionality. Sound/Speech interfaces have the advantage of not needing body-movement and of being activated using mechanisms that the users already have. With Speech or Sound therefore, the learning process of a system using this interface will be shortened. This interface may require attachments to boost accuracy, although the interaction remains the same whether with or without attachments. Gaze, EMG or Breath interfaces require longer training periods to achieve acceptable results.

**Aesthetics, Hygiene and User Acceptance.** User acceptance is related to aesthetics and hygiene which can limit the use of the technology. BCI is the interface that has the most invasive approach, requiring the placing of a set of electrodes in the user’s head or even the placing of devices on top of the skull, and therefore suffering from a lack of aesthetics or even hygiene if we consider more invasive approaches. Breath interfaces also suffer in hygiene and aesthetics on account of them requiring many pipes linking the machine to the user, and working by direct contact with the user’s mouth, creating some hygiene issues and if thorough cleaning is disregarded there is even the possibility of bacterial infections. Non-intrusive interfaces have advantage in
this evaluation topic since aesthetics and hygiene can be easily improved. Switch interfaces can be set where
the user wants or needs and activated when he/she needs, bypassing any hygiene problems. Therefore these
simple interfaces can provide high user acceptance. Speech/Sound and Motion Tracking interfaces have the
advantage of not requiring direct user contact and can easily be detected remotely so they provide users with
freedom of movement and minimal hygiene issues.

**Mobility Adequacy.** Many interfaces are prepared to work only in a type of controlled environment and
situation, which can reduce the overall system functionality since situations and environments can change
daily. The non-intrusive interfaces, which don’t need any attachments, are generally configured to work in
a fixed location and when their placement changes some adjustments may have to be made to achieve the
same level of results as before. Sound/Speech interfaces can suffer when in noisy environments or when
there are multiple sound sources emitting at the same time, which is a situation that is relatively frequent
when you consider the lifestyle in a house or in any outdoor scenario. Breath and BCI interfaces are not
suited to a mobile scenario, as they require a stable environment and specialised, generally large, equipment
to work, which rules out, at least for now, a mobile approach. EMG as Switches can be adapted to the user
depending on his or her needs, and the equipment that they need is small and portable. EMG uses electrodes
attached to the user, monitoring the user. Since they are attached, the user can move their body freely and
change environment without affecting the interface. Switch interfaces are the most adaptable to changes in
environment since they can be placed wherever the user needs or where he/she can achieve better results.
Their portability makes these interfaces suitable to space/user changes.

Taking into account the aimed user group, and considering the evaluation topics above, the interfaces
that may provide better results and be adapted to the users and to the users environment are clearly: the
Switches due to their high portability, user acceptance, ease of use and robustness to environment changes;
Sound/Speech due to the high information volume, user acceptance, and ease of use; Gaze/Motion Tracking
interfaces are adequate to a constant environment and are easy to use, presenting a high user acceptance
and information volume. When in a mobile or shifting environment, the EMG and choices are the most
adequate since they can easily be adapted and controlled using the same mechanisms as before. Table 2.1
gives a resumed looked over the assessed criteria in each technology.

<table>
<thead>
<tr>
<th>Adequated User Range</th>
<th>Dimensionality and Input Speed</th>
<th>Accuracy, Robustness and Repeatability</th>
<th>Ease of use</th>
<th>Aesthetics, Hygiene and Mobility Adequacy</th>
<th>Mobility Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Switches and Pointers</td>
<td>All</td>
<td><img src="symbol_1" alt="" /></td>
<td><img src="symbol_2" alt="" /></td>
<td><img src="symbol_3" alt="" /></td>
<td><img src="symbol_4" alt="" /></td>
</tr>
<tr>
<td>Sound/Speech</td>
<td>All</td>
<td><img src="symbol_1" alt="" /></td>
<td><img src="symbol_2" alt="" /></td>
<td><img src="symbol_3" alt="" /></td>
<td><img src="symbol_4" alt="" /></td>
</tr>
<tr>
<td>Gaze and Motion Tracking</td>
<td>LT</td>
<td><img src="symbol_1" alt="" /></td>
<td><img src="symbol_2" alt="" /></td>
<td><img src="symbol_3" alt="" /></td>
<td><img src="symbol_4" alt="" /></td>
</tr>
<tr>
<td>Breath Interfaces</td>
<td>HT</td>
<td>![symbol_5]</td>
<td>![symbol_6]</td>
<td>![symbol_7]</td>
<td>![symbol_8]</td>
</tr>
<tr>
<td>Brain-Computer Interfaces</td>
<td>HT</td>
<td>![symbol_5]</td>
<td>![symbol_6]</td>
<td>![symbol_7]</td>
<td>![symbol_8]</td>
</tr>
<tr>
<td>Electromyographic Interfaces</td>
<td>HT, LT</td>
<td>![symbol_5]</td>
<td>![symbol_6]</td>
<td>![symbol_7]</td>
<td>![symbol_8]</td>
</tr>
</tbody>
</table>

Table 2.1: Comparative Human Computer Interfaces Table.
2.3. Environmental Control for Tetraplegics

Physical constraints in Tetraplegics reduce the means with which they interact with the world, not only because many of the device control mechanisms aren’t suited to their capacities but also because their movement disabilities make it impossible to reach the desired devices. In a house, even for able-bodied persons, many control mechanisms are simply difficult to operate or reach a situation which is aggravated with motor-disabled users. A tetraplegic person in a house is left in the care of a caregiver without any control over his or her belongings, remotely operating his or her house. This situation represents a high workload for the caregiver, although, most of the time, their tasks are simple, e.g. turning the TV on. To increase user control over their environment, increasing their autonomy and reducing caregiver dependence, many approaches have been tried, although with some limitations or constraints. Below there is an explanation and discussion of the most relevant ones.

2.3.1. Robots for Environment Control

There are several robots whose main function is to give movement assistance. One example is the WHERE robot [49] [32] developed by KAIST for picking up a person, lifting him/her and transporting them to a desired location. These robots can decrease the caregiver workload drastically by decreasing the number of caregivers needed to lift or transport a disabled person in their daily routine. Bien et al.[6] transformed the Wheelchair into an adaptable device that could use various interfaces, adapting to the user’s needs. The system was called Kares II(KAIST Rehabilitation Engineering Service System II) and was developed at the Korean Advanced Institute of Science and Technology. There were 2 main modules in this project, the Robotic Arm and the Human-Robot Interaction (HRI). The Robotic Arm, constructed by the researchers, had 6 Degrees of Freedom with revolving joints and its main goal was to provide a way for the user to perform tasks far away from them. This robotic arm has 2 special functions, one of which is compliance with the user, to increase the safety aspect because during task execution it can have contact with the user. The other is the visual servoing that is incorporated in the arm to detect and locate objects and for analysing the user’s facial expression, like “opening the mouth” to issue a command that triggers the arm to give a glass of water [30].

The HRI had 3 interfaces available, each for one type of motor disability: Eye-Mouse, EMG and a Haptic Suit interface. The Eye-Mouse interface was designed for people with severe motor disability such as C4 injury. This technique uses a wearable device with a CCD camera that monitors the eye pupil and a magnetic position sensor that measures head movement.

The EMG technique was used for users that could make use of their shoulder or neck muscles for controlling the robotic arm or the wheelchair since the robotic arm was used as an extension of the user arm. There are 8 basic motions that can be detected by the shoulders to enhanced the recognition rate and make the system more adaptable the researchers used fuzzy learning algorithms. The Haptic Suit was a force sensitive resistor (FSR) that was used to detect the human motion. In this project just 2 parts of the body had been taken in account: Head and Shoulders.
VAHM stands for Autonomous Vehicle for people with Motor Disabilities in French and is a project that started in 1989 by G. Bourhis et al. [20]. This project aims to the creation of an intelligent wheelchair and was considered a pioneer in this field on those days. This project equipped a wheelchair with a set of ultrasonic sensors to navigate avoiding collision with the environment and to remove some workload from the user. There are 2 modes of navigation, the first is the Autonomous Motion Mode which makes use of all the system potentialities using a space segmentation technique and requires perfect knowledge of the robot moving area. The other mode is a Semi-Autonomous Mode that captures the users controls, like the joystick, and uses them to give a directional input to the wheelchair. To avoid injuries and increase the system safety that system uses a basic framework that avoids dangers like walls and obstacles removing some of the users navigating workload.

There are a number of mobile robotic arms available in the market as the technology gets ever smaller, making them useful for people that need to perform tasks at a reasonable distance from where they are but due to some injury or disease that have reduced their mobility to a short distance may see advantages of being capable of doing such tasks again with these assistive arms. The user interaction is made through a GUI that make interface with him/her by mainly a chin- or hand- controlled joystick, speech recognition, breath sensors and mechanical switches.

2.3.2. Smart Homes for the Motor-Impaired

There has been some development in the assistive smart homes, some examples of the early projects are the CASA project [47], HS-ADEPT [22], SmartBo project [18], Portsmouth Smart Homes and the Robotic Room [54] using different approaches to help the elderly or handicapped. Poulson et al. [47] developed the CASA (Concept of Automation and Services for people with Special Needs) was a two year project that started in 1994, it looked not just at standalone intelligent home technologies but also at how to integrate the hardware with outside support services. The emphasis of the project was on using home systems to monitor the home, mainly for safety reasons, by linking them to a remote service center. The project built demonstration systems in Spain and Portugal. The actual functionality in each flat varied, but commonly included gas and intruder alarms and appliance monitoring systems. The TV was used to provide an interface to the system and as an adapted remote control unit. In addition there were some limited control applications provided including remote door release, control of lighting, and an automatic night light. The project developed a number of prototype products based on the European Home Systems (HS) standard using the home power line as communication medium. There were some problems with the reliability of the prototype equipment but the project concluded that there was a significant role for home automation in the care and support of elderly people in their own homes, and that in the majority of cases, monitoring safety and security would be more appropriate than control applications.

Hammond et al. [22] developed HS-ADEPT in the same time span as that CASA project, addressing the challenge of developing interfaces for users of smart house technologies with severe physical disabilities who might be able to, for example, operate only a single switch. The system was based on bi-directional infra-red communications from the portable controller to a “gateway” to the EHS-Bus located in each room in the house.
This “gateway” held a model of all the home system devices that it needed to control in that room, as well as any others that could possibly be controlled from that room. Researcher concluded that users do not want to live in an automated home, they want control over their environment and that automation can be used to reduce the overall effort required in control.

2.3.2.1. ASPICE

One important project was the ASPICE project whose aim was “the development of a system that aids the neuromotor disabled to improve or recover their mobility (directly or by emulation)” [2] and had the goal of achieving maximum independence from the caregivers for its users. Preliminary results showed an increase in user independence although without substituting caregiver work. This project was separated into 5 modules: Input Devices, Feedback, Robotics, Domotics and Core Operations (Fig. 2.10).

The main interface used was a Brain-Computer Interface (BCI) due to its universal application in the target user group, but others could be used.

The system used a graphical user interface (GUI) as the user control unit which could control a AIBO robot [11] that was used to send voice messages to the caretaker such as “give me water” using it as a messenger. The GUI was also used as the feedback module. Therefore this module was the centerpiece of the interaction user-system, making the bridge between the system and the user. Since the system had a modular architecture they argued this module could use a low power computer such as a palmtop PC, thus not being a burden to the user and enhancing mobility.

The main home appliances are integrated using a set of specialised actuators. For the interactions, all over the house, this project used 3 technologies: X10; IRRC; Wifi; depending on the device type.

2.3.2.2. Intelligent Sweet Home

The Intelligent Sweet Home project is an ongoing project developed by the Korea Advanced Institute of Science and Technology (KAIST) for assisting the elderly and the handicapped that is based on 3 modules: Bed Robot System, Soft Remote Control, Network System.

![Figure 2.10: ASPICE Graphical User Interface](image-url)
The **Bed Robot System**, can be separated into 3 sub-modules: automatic bed, robotic arm and a motion capture device. The automatic bed is a bed that can be adapted to its users' needs and has 3 basic movements, which can move the basic body joints, namely the knee joint, pelvis joint, and spine joint, so preventing health problems such as bed sores, and giving more comfort and independence to the user. For the robotic arm the researchers used the MANUS robotic arm to extend the range of the user’s actions around their beds. The MANUS arm is configured for object transport functions with a limited work range that includes a table, a bookshelf, and a massage-hanger around the bed. These functions rely on are done by finding the correct position of the items, which are marked for recognition. This is done by a motion capture device which gives feedback to the system of the accurate position of the desired objects. The **Soft Remote Control System** [17] was responsible for recognizing the user's intentions and transforming them into commands for the system. This Control System recognizes predefined gestures like hand-motion and hand-posture using 3 CCD cameras by extracting the moving region [16] and adjusting the recognition to the user using Hidden Markov Models [48]. Voice control was another possibility in this system which used simple commands, with a maximum of 2 words, to start commands. Remote Control is used to control the surrounding environment by simply pointing at the desired device, whose position is stored in the system, to select it and using voice or gesture commands the user can confirm the device and give a set of commands to that device. There are 2 modes of control environment: Simple Mode and Extended Mode. The Simple Mode is used for on-off commands and the Extended Mode has a task menu that is displayed on a TV screen, and makes it possible for the user to choose more complex actions.

In the case of gestures, it had a recognition rate of 95% in the recognition of an object which was pointed at, and an overall rate of 80.7% to 95.7% which was similar to that of a regular remote control using infrared waves. The success rate depends directly on illumination, and normal movements like resting the hand can be interpreted by the system as a command. For voice recognition there were also some problems in noisy environments.

### 2.3.3. Evaluation Criteria

Each system has particular characteristics which are aimed at a specific user group. To make any kind of comparison a common set of evaluation topics had to be created, giving a clear background over future choices in assistive technologies. The evaluation topics were chosen by their capacity to segment the target

![Figure 2.11: Soft Remote Control.](image-url)
user population in an intelligible and objective way. The topics chosen were:

**Potential user range.** All assistive technology is aimed at a user group and its specific needs, limitations and capacities. This relation user/modality sometimes makes it impossible for other users to use the same approach. Taking into account each user’s limitations, severe in the tetraplegic’s case, not every assistive technology can provide the same kind of results for every type of user, thereby reducing its application. By characterising the requirements needed to operate a certain modality/assistive technology and mapping them with each user group’s motor and sensory abilities an assessment can be made that points out the potential user groups that could use the technology.

**Performance, Accuracy, Robustness.** This set of topics tries to characterise the reliability of the approach by explaining the situations that could provoke failure or error. With this set the consistency and maturity state of the system can be assessed. Performance specifies the system response to stimulus while Accuracy is concerned with the quality of the response and Robustness specifies the system failures or errors.

**User Adequacy.** All systems are used or tried by users and a set of arrangements have to be made, either physical preparations such as connecting the system components or pre-training users on how to use the system to achieve the expected performance. All these preparations affect system utilization. In all systems there are required actions, independent of the modalities used, which have a difficulty level. This topic defines the user experience in the system by characterising User Acceptance, Ease of Use and Aesthetics.

**Mobility Adequacy.** Many technologies are created to work in a specific environment and when the environment changes problems arise and results decrease. This topic characterises the ability that the system has to adapt to environments other than the environment which it was conceived for.

**Ease of Installation and Cost.** Systems may require specific types of hardware/software to operate which have to be acquired. By characterising the cost inherent to the use of the technology and specifying the availability and installation preparations that the system needs we can assess the economical barriers and the applicability of the solution.

### 2.3.4. Discussion

Despite increasing development in this field, the robot unit still needs high-rate low-level commands to be operated in a reliable way, which not every accessible interface can give, i.e. BCI. The main reason that these solutions don’t have a significant market share is the low and limited functionality and the high costs relative to other solutions like Nurses and Caregivers. All of these projects represent a solution to one part of the problem, giving more autonomy to the user, but all have these same disadvantages:

**Potential users range.** Robotic approaches can extend the user action range which is an advantage for people with movement difficulties such as Tetraplegic users. The type of users that can benefit from these systems is defined by the control interfaces used.

**Performance, Accuracy, Robustness.** These systems are slow but relatively accurate. Generally they are supported by precise environment mapping algorithms which increase the accuracy but have the disadvantage, in a changing environment, of system failure or error. Recovery from such errors can require
When the objective is constructing a system that aims to give control over home environments for a user with severe motor constraints (Tetraplegia), the system has stronger requirements than others that are aimed at able-bodied users. Reliability and system adequacy to the user/environment factor is essential to avoid con-

2.4. Overall Discussion

When the objective is constructing a system that aims to give control over home environments for a user with severe motor constraints (Tetraplegia), the system has stronger requirements than others that are aimed at able-bodied users. Reliability and system adequacy to the user/environment factor is essential to avoid con-
stant specialised intervention, such as ISH, ASPICE and robotic approaches to improve user independence. With this objective in mind, inputs have to be chosen. Speech/Sound interfaces can be used by almost all tetraplegics carrying a great amount of information while being available in every environment although its quality can decrease in some situations. Switches and Pointers can be a excellent choice for selection tasks due to their residual cost and environment/user adaptability. Gaze and Motion tracking interfaces can be used for spatial search tasks although user constraints frequently degrade the information quality. However if the user’s lack of motor coordination, which is directed but inconstant, could be used, it would represent a great choice in terms of graphical interface navigation. BCI approaches are still present with low functionality while carrying high costs so are only valuable with the most severe tetraplegics.

Robotic approaches are aimed at increasing user mobility and range and are the mechanisms used for appliance operations such as tetraplegics can perform. Most of these systems ease mobility by closing the gap between device and user but avoid more complicated operations such as activating/deactivating the appliance. The low functionality, lack of adaptability and the high cost condemn these approaches.

Of all Homes for Tetraplegics the ASPICE had wider user reach simply by using various input devices that could adapt to specific situations or users’ motor skills. Having gestures as the main input mode for tetraplegics could be argued as a bad choice due to their lack of motor control in their hands and arms (found in every tetraplegic), and fine gestures like in ISH can only be used with paraplegic users. The lack of comparative or performance results in these studies is astonishing implying that no or few user studies were done to prove system value in the target user group. A set of evaluation metrics has to be formed and decided upon to standardise these systems and make possible future comparisons.
Home Control for Tetraplegics

An Approach is a set of ideas or actions intended to deal with a problem or situation. Ideally its aims at finding a better way to solve a problem or limitation, delivering more sophisticated, more reliable and faster solutions to problems everyday. The Accessibilities research area is an area that has to deal with complex problems simply because limitations come from two fronts: Users and Technology; the users that we deal with are users that have limitations and constraints at a higher level than normal individuals. This fact affects the way that a user interacts with technology which, usually, isn’t designed to cope with their needs. The only solutions that may help these people require a successful assessment of the user profile and adaptation of the technology to the user’s needs. Nevertheless, that doesn’t necessarily mean that solutions found in this area can only suit impaired people. On the contrary these technologies can ease an able-bodied user’s life and relieve them of some of their workload, the most common example being the remote control, which can help both populations.

Having these assumptions in mind, the target user group must be thoroughly analysed and then existing solutions or technologies must be assessed. Then the advantages and disadvantages will lead us to a path that must be explored and tested with users, which may validate or invalidate the solution proposed.

3.1. Preliminary studies

Although technological solutions have been surveyed, the data provided isn’t sufficient to assess its functionality and usage in the target scenarios. Flaws and advantages can be assessed by these studies but validations and other problems must be taken into account to create a user-centred prototype. The assessment in this area has a higher importance that in others because motor-disabled users are difficult to reach and because their motor-skills vary from person-to-person.
3.1.1. User-Centered Design

User-centred design is a design philosophy and a process in which the needs, wants and limitations of the end user of an interface are given extensive attention at each stage of the design process. It can be characterised as a multi-stage problem-solving process that not only requires designers to analyse and foresee how users are likely to use an interface, but also to test the validity of their assumptions with regards to user behaviour in real world tests with actual users. Such testing is necessary as it is often very difficult for the designers of an interface to understand intuitively what a first-time user of their design experiences.

The difference from other interface design philosophies is that user-centred design tries to optimise the user interface around how people can, want, or need to work, rather than forcing the users to change how they use a system to the way the system must be used, creating further difficulties instead of helping. A Norman and Draper study, one of the first that were concerned with this philosophy, in their book [42] used the term "user-centred design" to describe design based on the needs of the user, leaving aside what they considered to be secondary issues like aesthetics. Clearly technology has evolved since those days, and aesthetics and usability factors have grown in importance since then.

Most studies of this kind starts with the user’s characterisation. At first researchers gather a number of people that they think could represent a segment of the population. The group must be large enough to support their assumptions. However when considering tetraplegics, an increased number of challenges is presented at every step of the process. They are users with health issues and limited motor-skills, which affects the way they communicate with the world. Their communication with the outside world is done through a caregiver who gives support to every activity in their daily life. The availability of the caregiver combined with the problems inherent to the target population creates unusual difficulties in reaching people with Tetraplegia, and some user studies are considered as Use-Cases which means that the researchers base and test their findings on a reduced amount of users and therefore their results can’t be generalised to the whole population segment.

In this stage of this dissertation only one tetraplegic, PF, was taken into account in the user analysis done, so the user study has the classification of a use-case. Although this limited the population reach of this research project, it contributed by creating a closer link between user and researcher, involving the user deeply in every stage of development.

3.1.2. User and Task Analysis

User Analysis is the process by which the user’s constraints and potentialities are specified. It is a subjective area that requires time and precise methods to characterise not only the user’s profile but also the current panorama that they are surrounded.

Initially this stage started with questionnaires and to specify the user’s motor-skills and interaction models regarding home appliances in his house, see Annex A1.1. The user used a laptop and was familiar with Windows interfaces which could be controlled using an arm stick to select the keyboard keys. Arm movements were used to operate the touchpad. The appliances in his house were controlled via caregiver. He issued
commands and the caregiver performed them, meaning that he did not have any control over the surrounding environment. He had an electrical wheelchair that he could control through arm movements using a joystick, which was the only situation that he controlled entirely. There was also an electric bed which could be controlled by a set of buttons embedded in it. However, the user didn’t have enough dexterity or arm strength to perform these actions. The more frequent tasks, done 3 or more times a day, were operations with the TV, Sound System, Mobile Phone and Computer. With the exception of the mobile phone and computer, all other devices were operated via the caregiver.

In terms of place, all actions performed were done in his bedroom where he spent 8-12 hours daily, in bed, having to change body position, again through caregiver help, every 2-4 hours to avoid further health problems. Occasionally he was in an electric wheelchair, though this was restricted to 8 hours maximum, 2-3 times a week, which required a rest period of 1-2 days to recover muscle strength. These were the only existent time restrictions.

The user demonstrated interest in operating their surrounding appliances remotely without caregiver assistance. The most important devices were: the TV, bed, lights, blinds, door, heating system and sound system. By being able to control these devices, the user could perform most of his daily tasks, relieving some of the caregiver workload and increasing his self-esteem and autonomy. In terms of tasks with these devices the most important are the activation and deactivation tasks, although other options may be important, while not essential.

3.1.3. User Interaction in Computer Access

The most common GUI\(^1\) are based on WIMP\(^2\) metaphors, the first system of which was Xerox PARC in the 1970’s, though it gained expression when the Macintosh used it in their system in 1984. These interfaces use simple physical and mental models, that even children can master, which makes them simple and easy to use. In these interfaces not only Pointing was embedded, but also the “Click” as a means of expressing intention by the user and with it many devices flourished, some of which are still in use today. Due to the close relationship between devices and GUI’s, every pointing device, nowadays, must be evaluated based on the WIMP GUI’s.

Limitations that users may have or characteristics that make it impossible for users to conveniently use the same modalities in a GUI, represent a decrease of input/output information and, consequentially, in performance. By tackling these ambiguities, using mechanisms or substituting/conjugating modalities, it’s possible to enhance or achieve the same level of performance as that achieved by people without modality limitations. This section makes reference to some of these approaches, starting with GUI mechanisms, followed by unimodal approaches and finishing with Multimodal approaches.

Having proved the importance of these interfaces for computer access, we must then analyse the existing solutions or approaches that could enhance the users interaction, either by interaction mechanisms or by GUI enhancers.

---

\(^1\) GUI stands for Graphical User Interface which are interfaces based on visual feedback.

\(^2\) WIMP is a concept composed of a set of elements: Windows, Icons, Menus and Pointing. This concept is used to built graphical interfaces.
3.1.3.1. GUI Mechanisms to Enhance Accuracy and Performance

GUIs are composed of components, selectable icons and menus, which may differ in size, form, distance, etc., metrics that affect pointing performance. Some approaches use these factors to change components localization and size in a way so that the cursor movement and selection is reduced to a minimum. Blanch et al. [7] proposed a separation between visual space and motor space changing the control-to-display ratio according to cursor distance to nearby targets, making targets fast to reach with variable velocity.

There are several magnifying techniques that use the cursor position to magnify the surrounding area. Microsoft has one solution included in their operating systems. These systems may be aimed at people with visual impairments but, conjugating with Blanch et al. approaches, other kind of systems can be developed (see next section).

When people using an input device have a low level of accuracy, other approaches can be followed. Trewin et al. [57] developed a mouse stabilization approach that suppresses mouse slipping when clicking for persons with motor or mental disabilities. Results showed a 25% reduction in target selection and 75% decrease in target misses, and this improvement also had high user satisfaction.

Another approach consists of the construction of GUIs that adapt to the user’s capacities. Gajos et al. [21] developed a set of systems to assess user abilities; SUPPLE identifies a user’s capacities, ARNAULD captures user preferences and the SUPPLE++ system transforms the interface according to the data provided by the other two systems, Figure 3.1. Tests with motor-disabled people showed a 24% decrease in task times and 73% fewer errors. There was a strong preference for SUPPLE++ interfaces over the manufacturer defaults, and users found them more efficient, easier to use, and much less physically tiring.

![Figure 3.1](image)

Figure 3.1: (a) The default interface for a print dialog. (b) A user interface for the print dialog automatically generated for a user with impaired dexterity based on a model of her actual motor capabilities.

Other approaches worked on the selection mechanism problem. Crossing interfaces, Figure 3.2, use the
same kind of components as the Clicking interfaces but make the selection action using only cursor movement. To make a selection, the cursor must cross over the desired component and then return. Results show promising results, particularly for people with motor disabilities or who are unable to perform a Clicking action. Although able-bodied users’ throughput was 25% inferior with goal crossing, compared to area crossing, the opposite was true for motor-disabled people, with an increase of 19% in the throughput [62]. Error rates were higher for goal crossing than for area pointing under a strict definition of crossing errors (6.23% vs. 1.94%). Subjective results indicated a preference for goal crossing among motor disabled users.

Figure 3.2: (a) Users with motor impairments often have difficulty acquiring area targets (b) In goal crossing, the need for acquiring a confined area and clicking is removed; goals must only be passed.

3.1.4. Domotics

To cope with the proposed objectives, which are based on low cost and high functionality, existing domotic technologies must be evaluated and discussed. There are some technologies to Home Automation [59] but the relevant ones nowadays, taking in account the number and variety of devices for sale on the market, are the bus based communication systems that use a power line or a separate cable wire, some examples are X10, EIB/KNX and LonWorks, and the wireless systems such as 802.11x IEEE standards. In Fig. 3.3(a) and Fig. 3.3(b) [45] we can see a timeline and an outline of the existing home technologies. The most common and most easily available on the market are the X10 and the EIB/KNX. The first carries low functionality and low cost, whereas the EIB/KNX which is costly but has higher functionality.

Figure 3.3: Domotic Overview.

EIB and KNX
EIB (European Installation Bus) is an open standard widely used in Europe. EIB is available for power line, signal cable and radio. The single cable version is currently the most commonly used in smart homes. KNX is a new standard resulting from an amalgamation of three European bus Standards, with EIB being one of them. KNX is expected to replace EIB in the near future. KNX fully complies to the EN 50090 series, the European Standard for Home and Building Electronic Systems.

Both of these technologies require highly-specialised configuration and specific modules. Although they are open standards, specific formation and certification must be taken in order to work with the technology. This creates a barrier because all the platforms used to configure and operate the systems are controlled by a specific association, Konnex Association, which allied with the high costs of the equipment, configuration platforms and required formations carry a very higher cost.

X10

X10 is an international and open industry standard for communication among electronic devices used for home automation. It uses the power line wiring for signalling and control, where the signals involve brief radio frequency bursts representing digital information. A wireless radio based protocol transport is also defined but it’s not an original standard of the protocol. X10 was developed in 1975 by Pico Electronics of Glenrothes, Scotland, in order to allow remote control of home devices and appliances. It was the first domotic technology and remains the most widely available because of its low price and ease of use, but has some problems that limit its functionality which are the lack of accuracy due to some electrical anomalies, the reduced command rate, its high transmission time, and the small amount of information that can be sent or received. The X10 technology’s greatest limitation is the fact that it can’t ascertain the device’s current state; there is no way to find this out automatically, so the user has to visually check the state of the device. Although it has several limitations, this technology can easily fit the essential user needs, but in systems that have high or moderate demands it may not be the best choice.

Domotic Discussion

The analysis of the characteristics of each technology pointed out the X10 technology due to its low cost and sufficient functionality to cover the most common and essential needs of its users, i.e. activating and deactivating electrical devices, while at the same time being very easy to install due to its transmission channel being through the existing electrical wiring of the house. Its range from the source to the desired device is 80m which is more than sufficient to cover a medium house or division. With this technology there are some constraints regarding the number of devices that it can cover, it being the most limited of all surveyed one with a limit of only 256 devices, but, nevertheless, this is more than enough to control a house.

3.1.5. Current Scenario

In every study an assessment must be make taking into account the users and the scenarios that they are found in. The objective of a scenario is to give an overview of the interaction existent between user and environment, especially when an operation is performed on the house appliances. Using a scenario as an
example we can give an outline of the existing interaction in a house.

"John was in bed at home, lying down, leaning to the right when he remembered that his favourite TV show was going to start in 5 minutes. He then shouted to his caregiver until he/she answered and arrived in his room. Then he asked the caregiver to turn on the TV and select the channel SOC which was his favourite channel. Since the temperature in the room was high he then asked the caregiver to turn the heating off. He then felt thirsty and asked for a glass of water which implied bringing the water and supporting the cup with a straw while John was drinking."

3.1.6. Preliminary Requirements

After the questionnaire analysis a set of requirements were assessed. They are divided into Functional Requirements and Non-Functional Requirements.

Functional Requirements characterise the functions or the behaviour that the system must have to achieve its goals and Non-Functional Requirements characterise the constraints or qualities in a system. Below there is a description of some of the found requirements. These subsets were discerned via questionnaires, interviews and observations.

**Functional Requirements**

- The system must be capable of controlling any electrical appliance in a house by at least activating or deactivating it.
- The system must provide a way to remotely control home appliances, while maintaining their original interfaces.
- The system must provide ways to control the devices remotely independently of the user’s position.
- The brightness of the Lights and the level of the Blinds must have the function to be raised or lowered.
- Every modality can be used simultaneously.
- Every operation can be sent by each modality.

**Non-Functional Requirements**

- When a command isn’t recognized the maximum number of retries is 2, meaning that no later than at the third attempt the command must be accepted.
- The time between a successful command and the appliance action must be less than 5 seconds.
- When an incorrect action occurs the user must be able to correct it with no more than 2 operations.

3.2. Environmental Control for Tetraplegics

Environmental Control is concerned with almost all the things that surround us, so has a very wide scope. In this dissertation we will only concern ourselves with one subset, which is the home environment that surrounds...
tetraplegics in their daily life. With the objective of increasing the user’s autonomy and action range, many scenarios were studied and objectives were found that lead this dissertation. By taking into account the users’ characteristics, daily life and the environments that they found themselves in, we could have a glimpse of what users really need and lack.

The analysis which was done revealed that the users, although restricted because of their physical disabilities, could perform simple, inconstant and inaccurate gestures but with a relatively stable direction when using the arm. These potentialities aren’t made use of due to the lack of mechanisms that could exploit them, instead leaving the caregiver to do almost all the tasks. Regarding home appliances, most of the tasks done were to turn appliances on or off. Occasionally, depending on the appliance, other actions were done such as increasing or decreasing the volume or changing the TV channel.

The next sections will explain what we expected the system to perform or what the objectives were and the main directives that it was based on.

### 3.2.1. Objectives

Every system starts with a set of objectives which are defined by taking into account the user’s needs and what the researchers think that they could improve, also known as added value.

Returning to the user profile, they are tetraplegics with low Tetraplegia that have reduced arm movements and intact muscle movements above the shoulders, meaning speech is intact. They live at home and are monitored and helped by a caregiver. Most of the time they are in bed and have access to, and control of, mobile phones and portable computers. Regarding their technological knowledge, they are expert users of Windows interfaces. There are some economical problems with these users, since they live at home and have to pay for caregiver help, a situation to which the government doesn’t give enough economic support, leading to some economic constraints.

The set of objectives proposed were:

- Use the gestures to select actions in an accurate way.
- Use the Speech modality to complement or enhance task performance or accuracy.
- Create a system that not only can address specific situations but also can be adapted to work in different home environments.
- The modalities used should be such as the user desires, meaning that any combination of modalities must be possible. This objective extends the adaptability of the system because when situations change, such as when in bed or in the wheelchair, some modalities may become more inaccurate or difficult to use, so by enabling every modality to be used at the same time there is a wider adaptability.
- Create a system that could cope with most of the needs of the users while being robust, cheap and adaptable.

Although there were few objectives at this stage they are bold and pioneering in terms of systems for tetraplegic users.
3.2.2. Aimed Scenarios

Aimed Scenarios represent the ideas and objectives that we would aim for, guiding the system in a way so that the scenario could be possible. It's a contextualized representation of the expected potentialities of the system in an understandable story form. Ideally this description should be possible or have similarities when compared to the final scenario. However, many stages and tests lie between this stage and the final prototype, so choices can be abandoned or intensified depending on the results achieved.

"In the evening, John asked his caregiver to move him to the wheelchair. When in the chair John headed to the living room, as Rocky III was scheduled to appear on TV. He stopped by the sofa and said **TV On** which activated the appliance. Then he looked at the windows and said **Curtains Down** which closed the curtains. When the movie ended John went back to bed but forgot the TV was on, so then he said **turn the TV off**, but because there were 2 TV sets on, one in the bedroom and one in the living room, a visual feedback was shown and then he rephrased the instruction **turn the TV in the living room off**."

3.2.3. System Inputs

An Input device is a sensor which has monitors a characteristic and sends information to another system. This module is responsible for capturing the user’s intentions and translating that into valid and understandable information which will be used by the system to trigger actions. It represents the frontier between the system and the user which the user uses to express intention. To control the system, there are two control interfaces: Speech and Gestures. The choices are justified below and are based on the analysis of the user profile and the current technological state of the art.

**Speech**

Speech is a form of verbal communication based on sound and therefore it can be captured by any sound capture device using a direct input method. There are fixed and portable microphones which can be suited to the user’s needs and situations. Although some problems still exist in noisy environments due to its wide dimensionality and high data rate, we hope to decrease that by decreasing the recognisable dimensionality, reducing the grammar according to the system’s available options. By decreasing the dimensionality we can raise the accuracy and robustness of the recognition system, but the noise problem will still exist which can be detected by recognising the false-positives. The grammar must be specified according to the user’s vocabulary, the house hierarchy and available devices to achieve successful results.

**Gestures**

A gesture is a form of non-verbal communication made with one or many parts of the body. In the case of tetraplegics, gestures may not be the ideal form of communication since they have severe physical constraints which reduce their movements. However, they are able to perform simple movements with their arms, although inconstant and with reduced precision, a direction can be obtained from them. That direction can then be the
input for the system.

Gestures can be captured by control interfaces that can be fitted according to the user, i.e. if the user is experiencing some tiredness in the arms, he/she can then use the neck muscles to perform similar movements. This factor enhances the adaptability of this modality, although gesture capturing devices must be selected according to the movement types and body locations that will be captured.

In this stage, two control interfaces were analysed, the first using a touchpad and the second using a laser beam. The touchpad was chosen because users use this device for computer access and are familiar with its interaction mode. This device is a pointing device consisting of a specialised surface that can translate the motion and position of the user's fingers into a relative position on screen. It's very easy to use and doesn't require strength to operate. People with physical constraints use the touchpad instead of the mouse because of its low amplitude and low-strength movement requirements.

Laser pointers are useful pointing tools for pointing, generally in large displays. Their main advantages are portability, low-cost and a distinct near zero-length dot that can be easily recognised. A distance transition, from far to close to a display, is made naturally and is similar to a touchpen interaction. With this device we hope to achieve better results and provide an option for monitoring the system.

The large range that the laser pointer offers us, allied with its versatility, makes it easy to adapt to the user's body. A laser pointer can be fitted to any user body part due to its residual weight and dimensions. This represents an enormous advantage over other devices which require a higher level of adaptation.

### 3.2.4. Proposed Input Methods

Input Method characterises the way used to transform data into intelligible information, it usually specifies a Interaction, which is a kind of action that occurs as two or more objects have an effect upon one another. The idea of a two-way effect is essential in the concept of interaction, as opposed to a one-way causal effect. An input method is a way to produce meaningful information from the existing interaction between the user and the system.

**Matrix Interaction** The input method proposed was based on a 3x3 matrix concept which means that there were 9 possible movements, with each being cell associated to one of them, Fig. 3.4. The method works by dividing the visible image into an odd number of cells, both in columns and in rows, and by delegating to each cell a direction and a velocity. The direction of each cell is calculated based on the vector between it and the centre of the middle cell. By default it doesn’t have velocity or direction, and the centre of the cell that the cursor is included. Due to the large size of each cell, the influence of hand jitter is expected to be minimal even with high jitter. The middle cell assumes a stabilizing behaviour because it prevents the cursor from moving through it when it would otherwise be moving.

**Absolute Interaction** In contrast with Matrix Interaction, Absolute Interaction uses an absolute approach in which any movement in the input device is immediately reflected, this input method doesn’t use any movement filtering, directly reflecting the user’s movements.
Gesture Dimensionality

The gesture recognition algorithms are responsible for classifying the movements into valid gesture classes. To reduce errors and increase the success of this module, which is constrained by the user’s physical limitations, a set of classes were defined which correspond to 8 directions. This idea was based on a 3x3 matrix metaphor in which the centre cell corresponds to no movement and all the other cells define a direction. The vertical directions available are Up and Down, the horizontal directions are Left and Right and the other four directions represent the diagonal movements.

With this reduced set of valid options there might be a greater chance of success in recognizing the user’s intentions. However, by reducing the dimensionality we also reduce the information quality that can be gathered in the system that uses the gestures, delegating to them the responsibility of validating the outputs of the gestures.

3.3. Available Work Evaluation

An idea needs to be developed in order to test its characteristics, to see if there are already materials or pieces of material that could reduce the development time while producing the same output.

The proposed approach is based on a few basic assumptions which are the construction of a low cost system that can be used on day-to-day tasks requiring low maintenance while being easily adaptable to the environment and the user’s motor-skills and desires. Although it’s easy to say that we aim to built the “perfect system”, there is an inherent high complexity involved in developing it. Therefore, to reduce its complexity without reducing its functionality and complete it within the scheduled time, some parts of existing materials had to be used in order to accomplish our objective, both in functionality and on time.

This section gives an overview of the existing usable solutions found or, if the solutions aren’t able to provide the needed adaptability or personalization, the materials that must be developed.

3.3.1. Hardware Tools

Hardware is a general term that refers to the physical artifacts of a technology. They are used to capture or act over the environment, being the frontier between users and system. The Hardware tools needed for the implementation of the system are available at a low cost and are easy to install or configure.
Input Devices

To recognise user movements, two devices were used, one for detecting the laser dot and the other for detecting hand movements. The Laser recognition module used a regular Webcam for capturing the dot position, and a regular laser pointer.

For detecting hand movements, a set of input devices can be used to evaluate the user’s interaction, mainly the mouse and the touchpad. The mouse is a device commonly used by able-bodied users to access computer information, whereas touchpads and joysticks are commonly used by people with motor constraints.

Both devices are widely available and have minimal costs. Another advantage is that these devices are considered standard for computer access and therefore they are frequently embedded or offered with computers.

X10 Domotic devices

A domotic device acts as the bridge between the system and the device action by being connected and responsive to changes in the domotic network, which, using X10 technology, is the electrical wiring of the house. These domotic devices can be different, depending on their purpose. There are two X10 device classes, the actuators and the listeners. Below there is a brief explanation of these devices.

Actuators/Listeners and emitters

There are 2 main types of X10 devices, devices that send information and devices that receive information, the first type is responsible for the transmission of commands to the power line that reaches the desired devices, the receivers that read the commands and alter their state depending on the command.

With these two types of devices, and some imagination, this technology can create more complex and usable systems. By combining a receiver with a sender it’s possible to create a bi-directional terminal capable of giving information of its current state. However, it doubles the cost and due to network inconsistency could result in an error prone system.

3.3.2. Software Tools

For domotic control using X10 technology there are some options, which are the Active Home products the most complete⁴. However it isn’t open source and therefore is impossible to use or be adapted to other solutions. Taking that major limitation into account, it was decided that a library implementing the X10 protocol which is public and free to use, would be the best choice, although not the easiest or the fastest.

For recognising Gestures using a laser pointer, much research and material is available, which reduced the development time. Since the recognition for this device is based on image processing algorithms, an array of image filters can be used in order to accurately detect the laser dot. There are already libraries that offer these functionalities like OpenCV⁵ which is open source and therefore free to adapt if needed. Other Gesture recognition solutions using other pointing devices such as touch pad or mouse exist, but aren’t suited

---

⁴Available at http://www.activehomepro.com it’s a paid solution, without any library attached, that allows future works around this platform.

⁵Available at http://opencvlibrary.sourceforge.net
to our recognition algorithms, so an application was developed in order to experiment and test the proposed algorithms.

3.4. Preliminary Tests

Testing is the stage in which the researcher can validate the choices taken. However, in a system like this the test session must be separated by modules since it’s composed of many distinct modules which have different usages and functionalities.

Since the system was still in an early stage, this testing phase addressed only the input methods using only Gestures, using the touchpad and the laser pointer. By testing the most frequent input devices and a Laser Pointer, the user’s movements could be assessed and further analysis should point out the most problematic devices or situations that constrain the device or user, decreasing the movement quality.

These tests focused on the evaluation of pointing devices in spatial search tasks, using a set of measures to characterise accuracy and performance.

3.4.1. Motivation

Users control a system by means of its input devices. These devices represent the frontier between the user and the system and are an essential part of the interaction. Therefore this part of the system must be the first to be tested because all the interaction depends on it.

In the case of tetraplegics, the interaction has some constraints and regular input methods may not fit their needs. These users need a personalised evaluation that could point to the control interface or input method. Many studies have been done concerning tetraplegics, e.g. Tanimoto et al.[55] which described a measurement and evaluation system of the computer input ability of patients with tetraplegia which enabled the medical staff to choose the best input device for the tetraplegic for controlling a computer.

3.4.2. Related Work

Laser Pointers, due to being their being a freehand device, suffer from user hand jitter, Olsen and Nielsen [44] designed interaction techniques to avoid this problem, but its "On-Off" interaction still made it an error prone device due to frequent pointer misplacing. There are many studies that have tried to tackle these problems. Oh and Stuerlinger [43] designed a laser pointer controlled by computer, and results of their experiment showed a 40% error rate when selecting large, 40 pixel, targets. Myers et al. [40] evaluated laser pointers and found out that laser pointers performed the worst in comparison to other devices, even after the predictive filtering technique. With their Semantic Snarfing technique the results obtained were good even in comparison with a mouse. Their experiment suggested that direct devices performed better than indirect devices. Some relative or indirect interaction studies have been performed, such as Wilson and Phan [61] who used relative movement to control a motor actuated laser beam and discussed the differences between absolute and relative interaction.
Measures need to be assessed for inter-device comparison. There are two common measures in nearly all studies; the Movement Time (MT) which characterises movement speed, and Error Rate which characterises accuracy, usually specified as a percentage of wrong selections. Fitts’ law was the first to try to capture this human feature and ever since then his law has been perfected and studied [1].

Today there are international standards, the ISO 9241 ("Ergonomic design for office work with visual display terminals (VDTs)") Part 9 ("Requirements for non-keyboard input devices") standard [25] normalized pointing device evaluation. It establishes a point of comparison between studies, although it isn’t sufficient to capture all the complexity of a pointing movement. Mackenzie et al.[33] proposed a new set of measures for pointing device evaluation that better specifies movement and, due to its adaptability, has been used as a complement of other measures. Keates et al. [29] went further in his evaluation and proposed six new measures to specify fine-grained movements in motor-impaired computer users. These additional measures are useful not only for motor-impaired users but also for other users when the accuracy of the device or user is deficient, such as with natural hand jitter when using a laser pointer.

3.4.3. Users

Since there was only one tetraplegic person available for testing and the devices used may require fine-grained movements, we found that the value that could be taken by these tests was weak. Thus, in this early stage, the tests were done with able-bodied people to establish a baseline that could then be used for comparison in future tests. Although the tests weren’t done by tetraplegics, this first stage is expected to make the screening for major issues regarding the interaction with the devices. This reduced the iterations and workload for tetraplegics.

3.4.4. Participants

Eleven participants (10 males and 1 female), of average age 24.3 (SD=2.29), were chosen based on availability at a local investigation department. All participants were familiarized with WIMP GUI’s and the most regular input device used by them was the mouse.

3.4.5. Apparatus

A 1.7 GHz Centrino with a 15” screen laptop was used running Microsoft Windows XP Professional with a resolution of 1400x1050. The software was developed in C# .Net 2.0. The Windows Cursor speed was set to 7/10, the keyboard speed was 500 pixels/s and the Laser pointer was 170 pixels/s.

For detecting the Laser dot a Creative Live! Cam Vista IM using developed C++ software was used. The input devices used were: Mouse (600 dpi Mitsai MS-3662), Keyboard (QWERTY Portuguese Version Keyboard), Trackball (Logitech TrackMan Wheel), Touchpad (ALPS Programmable Touch Pad) and Laser Pointer (Regular Class IIIa Wavelength 635nm).
3.4.6. Procedure

The experiment was done in one session with five test groups, one group per device, one participant at a time. The tests in all groups had the same order but different nominal difficulty, (2, 3 and 4 respectively) all of these tests were the Multi-directional Tapping Tests of the ISO 9241-9[25]. The first test had 4 shapes and the others had 8 shapes and all were disposed in a circular layout. The target shapes had the form of a square and the diameter of the layout ranged from 150 to 330 pixels.

For using the Laser there were 2 modes. The Absolute mode does the direct mapping of the laser dot in screen coordinates. Joystick mode divides the viewable area into a 3x3 matrix. This mode moves the cursor in the direction of the cell in which the laser dot is, except for the centre cell which doesn't move the cursor and is used to stop the cursor. With this device the users were requested to have the Laser Pointer always On and if possible not to support their arm in any way.

At the end of the test session each individual completed a Device Experience Questionnaire and the ISO 9241-9 Device Assessment Questionnaire.

Each session started with a description of the tests and at the beginning of every test group, when the input device changed, a period of habituation started. When participants wanted to start the tests, they had to do a “Click” on the active shape, the only “Click” in any test group. To select the other shapes, the participants just had to move the cursor to the active shape, which was coloured red, and let it stay inside that shape 500 ms with a tolerance of 5 pixels of movement.

3.4.7. Design

In each session all 5 devices were used in the order enumerated in 4.3. For each device a group of tests were done, with each group having 3 tests. Each test had a set of shapes for the user to select coloured white, except for the active shape which was red. When the active shape was selected the opposite shape was activated clockwise, 3.5. The devices were placed one-by-one always on the right side of the laptop, except in the case of the touchpad which is embedded in the laptop. For the Laser, a white, vertical, A4 sized placard was set beside and backward of the computer with a distance of 40 cm, from the start of the laptop to the placard. In it 5 points were drawn, 1 representing the centre of the screen and 4 representing the screen limits.

Each test had a nominal difficulty 1 bit higher than the one previous to it, the tests started with 2 bits and ended with 4 bits.

There were: 11 Participants X (4 Devices + 2 Laser Interaction Modes) X 1 Test Group X (3 Shapes + 7 Shapes + 7 Shapes) = 1122 results
3.4.8. Discussion

In terms of Movement Times, the Mouse was the fastest of all devices followed closely by the Trackball and Touchpad. Keyboard was the slowest of all non-laser devices, as expected, since it is a button-controlled device. LMJ\textsuperscript{6} had a higher Acquisition Time than LMA\textsuperscript{7} in all difficulties, although, at low difficulties, Time is inferior to Click by 33-59%. In the Index of Difficulty (ID) 4, LMA is 15% faster than LMJ, suggesting that at low difficulties LMJ could represent a great improvement over LMA. The results of LMA for greater difficulties can be explained by the use of the table desk to support the hand and by the use of both hands to stabilise the pointer, creating also a relative method of interaction, invalidating that assumption, Figure A1.3. Further details can be found in Appendix A1.2.

The approach based on a Matrix metaphor could enhance the performance of a laser pointer device by 33-59% in comparison with a regular direct-based approach of the same device, while decreasing the jitter effect inherent to the users. In every device or interaction mode tested, the Acquisition Times and Selection Times were statistically significant. The performances of the Laser Interaction modes are greatly inferior, in comparison with other input devices, as much as 3-4 times in pointing-selection tasks.

The Keyboard had the worst performance and accuracy of the non-laser devices, as expected, due to the lack of control of the buttons. Since it is a button-controlled device, which works by On-Off, a short acceleration in the first movements was proposed and should increase accuracy when targets are close to the cursor. Trackball, Touchpad and Mouse had similar performances, except in task axis distance. Trackball and LMA were the devices that most crossed the Task Axis but the Trackball had a more precise path, closer to the Task Axis, and therefore a more precise movement.

LMA had interesting times for the highest difficulty tests, being faster than LMJ. This result can be explained by the use of hand support, by using both hands and by using the table as a movement stabilizer. The fatigue inherent to the lasers had two factors, referred to by users in the conversation after the tests and by the

\textsuperscript{6}LMJ stands for Laser Mouse Joystick due to it's Matrix interaction mode that simulates the interaction of a joystick

\textsuperscript{7}LMA stands for Laser Mouse Absolute because it uses the Absolute interaction mode, therefore every movement on the device is directly reflected to output.
ISO Assessment Tests, due to being the last device to test and requiring the pressure button to activate the laser. The advantages of these devices may come with their use at longer distances and due their portability. A Laser device without a pressure On-Off button could greatly decrease user fatigue but not user comfort. Users all felt uncomfortable about pointing beside the PC, so an approach of pointing directly at the screen should be studied.

Although it has a performance level 2 times below that of the mouse, LMJ has proven that it could enhance pointing precision drastically and thus make normal laser pointer interaction obsolete.

3.4.9. Post-test analysis

From the results observed in this test session and by comparing the interaction of the able-bodied user with the motor-skills of the tetraplegic, we decided that the interaction methods that were being tested imposed even greater restrictions on the motor-disabled users. This assumption was based on the high errors expressed by the laser pointers and the amount of precise movements needed to select a target. Although some difficulties were expected, they are strong enough to invalidate these interaction modes from the beginning.

The main reason for these results, although they showed a great increase in performance when using Matrix interaction in laser pointers, was the increased workload, tiredness and concentration needed to use the Absolute or Matrix interaction mode. In the Absolute mode users needed to make fine-grained movements to accurately reach their targets, at which all the users expressed difficulties. The difficulties varied between the input devices, but laser pointers, which are free-hand and direct input devices, had an increased difficulty due to the expected hand jitter. In the Matrix mode, laser interaction was made unnatural due to the lack of feedback. The users didn't know which cell they were pointing at, and therefore the adjustments made were done by trial and error. Although promising results in comparison with the Absolute mode were found, we need to be more demanding because of, taking into account the user profile, the inability of the users to perform fine-grained and complex movements.

However, the Matrix interaction mode clearly showed that the concept on which it was based could be valid and present acceptable results. With this same concept in mind, a new interaction mode, based on the same principles, was proposed.

3.4.10. Slash Interaction Mode

This mode used the 3x3 matrix concept to define the program interface and the gesture detection algorithms. The same reduced gesture dimensionality is used, although the concept of movement changes. In other interaction modes a movement was conceived as a set of positions that changed in time, creating a path that finally could be analysed. However, as seen on the LMJ which used the matrix mode, users have difficulty in realising the initial position or the place in which the movement is expected to start or end. Therefore another interaction was proposed that should deal with these problems without degrading the results of other interaction modes.

The studied user group has severe physical limitations. However they can perform simple, undirected
movements which don’t require a high degree of accuracy. An approach which doesn’t require them to aim at some target, either to start or finish a task, should have a better performance than others that require such efforts. There have already been some studies like Wobbrock and Gajos [62] which compared the area-pointing and goal-crossing interaction methods in GUIs and found that motor-disabled users could benefit from these systems. Our proposal simplifies even further the interaction in GUIs because it doesn’t require a specific area that must be selected and therefore the user isn’t forced to aim at that target.

Interaction that uses simple strokes or slashes, which are fast movements that have a direction that usually is constant during the movement, could provide an advantage since users can easily do these movements. The simple movement allied with the reduced gesture dimensionality could provide a reliable way to successfully detect the gestures of these users. In the Slash concept another concept is embedded which makes the interaction even simpler. That is the possibility to start and finish a gesture wherever the user wants, i.e. the user doesn’t have a start or finish target to impose restrictions. Working the Slash mode just requires a movement with a direction that, when inconsistent, delegates the analysis of the movement to more sophisticated recognition algorithms, Figure 3.6.

Although this interaction mode has some potentialities, it imposes some restrictions and a change of paradigm in the graphical user interfaces that will deal with this interaction mode. The usual WIMP interfaces are based on a few concepts, which are targets that are selected via area pointing mechanisms. However, the Slash mode isn’t suited to use in area pointing, which means that the interface that works with the Slash mode will use different metaphors such as the same basic concept of this interaction mode, the matrix concept.

Figure 3.6: Existing movement classification using in every Matrix-based interaction mode.
This chapter was based on Chapter 3, which defined our guidelines and concepts, to create a system combining several technologies in an intelligible and flexible way. The value of this system is expressed not only by the modalities or new interaction modes found, but also for the overall adequacy to the tetraplegics by enabling them to cope with the shortcomings present in their home environments, either caused by the users’ limited motor-skills or by the unfit surrounding environment. Indeed, with this system we aimed to enhance the users’ reach to their home appliances in terms of both quality and range.

The system developed was based on the principles of modularity, meaning that the functionalities were separated into modules, and flexibility, meaning that the architecture allowed the combination of the modules in a reliable and fast way maintaining the same level of performance and functionality. These factors made the corrections and additions that led to the final prototype possible and then easier.

4.1. System Overview

The system built covered many technological areas which had to be enclosed in particular modules. That fact contributed to an architecture with more levels than expected since when we use different technologies or systems for different purposes we must specify their features in a way so that they could be reused by other levels.
The system can be divided into 4 parts: Inputs, Outputs, Domotic Network and Core Operations, Figure 4.1. Below there is a brief explanation of each main module.

**Inputs**, is the module that deals with all the aspects of the interaction between the user and the system. In it there is a cursor tracking system that collects data and, when a movement is finished, transfers it to the recognition submodule. This submodule implements all the movement recognition algorithms tested. The input devices that use this module are a microphone, a laser beam and a touchpad.

**Outputs**, implements both graphical and auditory feedback. Although two distinct modalities are used for feedback, they are triggered in different stages. The auditory feedback is triggered when an action occurs, emulating speech to describe the action triggered. The graphical feedback gives a visual overview of the actions detected in real-time and the current state of the house appliances. It provides feedback from both gestures and speech commands, reporting to the user both possible or accepted actions at that time.

**Domotic Network** is the module responsible for the actions in the household appliances. It receives commands from the system and reflects them in real actions in the household. It enclosures all the physical layers of the system, being as it is, the module of the system that delivers all the functionality to the users.

**Core Operations**, is responsible for managing the data sent by the input devices and the feedback returned to the user. It represents the centrepiece of the system, constructing the graphical interface and analysing the data sent from every input device available.

The next section will explain in more detail the system modules, specifying their information flow.

### 4.2. Material Used

Although this system was developed to be adaptable to new input devices, only a few were tested. Their insertion in the system was done by means of different techniques. Below we will explain the materials that were used by each module.

**Inputs**
A regular Class IIIa Wavelength 635nm Laser Pointer was used for pointing. However, to recognize the laser dot an image capturing device was needed and therefore a Creative Live! Cam Vista IM was used which, using an image processing program, enabled the detection of the laser dot without the need for optical filters to be attached to the camera.

For detecting movements, some input devices were used in the preliminary tests, but just the touchpad survived due to availability and performance factors already discussed. The touchpad used was a regular touchpad of dimensions 6cm x 4.5cm.

Speech was the only input whose aim was not the emulation or detection of gestures. For this modality to be used, a sound capturing device was used with a speech recognition system that could detect the user’s speech. The device used was a regular microphone, Labtec Stereo 242.

The devices for this module and all other modules, except the Domotic network module, can be easily found or substituted by other devices that are similar. There are no specific requirements for the devices that must be used for the system to work, enabling the users to use the devices that they already own.

**Outputs**

There are no special requirements for the Output module. It uses any computer screen for graphical feedback and any sound set for auditory feedback. In this study both of these devices were embedded in the laptop used.

**Domotic Network**

There are 3 main types of devices; devices that send information, devices that receive information and devices that connect the system to the PC. The first type is responsible for the transmission of commands to the powerline to reach the receiver devices that read the commands and alter their state depending on the command. The PC interface can send or listen to commands in the powerline, storing some information if needed.

This was the most demanding module because it required specific hardware. The hardware used is standard in the X10 industry, and can be found in places that sell domotic hardware. A CM11A device was used as the bridge between the computer and the domotic network. It was used to listen to the electrical network and to send commands. This device is the central part of all domotic operations, as it is responsible for all operations in the network. This interface receives a command, which it then delivers, according to the X10 protocol, to the domotic network. This command is then read by the X10 actuators. The actuators used were LM12G and AM12G - lamp modules and appliance modules respectively - which listen for commands present in the power line of the house, answering the actions that have their identifiers.

**4.3. Multimodal Interface**

A Multimodal interface is an interface that groups together a set of modalities in a way so that they can be used to enhance or compensate for the shortcomings of each other. When we use these interfaces, it’s necessary
to have a in-depth knowledge of the limitations and potentialities of each modality, in order to be able to create the proper connections between modalities and to choose the modalities that may perform better in a specific context. The choice of using Multimodal interfaces in this system was a natural one on the basis of the user analysis which clearly showed that some modalities had serious deficiencies, especially the gesture-based ones, but that these could be minimized when also using speech-based inputs. This simple principle lead us to the development of a Multimodal interface that used Gestures, in a number of forms, and Speech. Both of these interfaces have some limitations either due to the quality of characteristics captured or due to the user’s limited motor-skills. In this section these interfaces will be explained by modules, with an overview given of how and what is included in the system.

4.3.1. Speech Module

The Speech Module captures sound from the system default microphone and transfers it to the system speech recognition engine. In this module a standard free-to-use system recognition engine from Microsoft was used, which was the Microsoft Speech Engine 3.1. This engine detected speech from the sound captured, although the speech that can be captured can be restricted by grammar\textsuperscript{1}. The grammar was built to have a user-centred approach, using the data collected by a set of tests which were done with a tetraplegic user; Section 5.2 makes reference to the conclusions taken from these tests. Further to the development of a suitable grammar, the tests also determined the meaning of the commands and allowed a distinction which enabled the separation of the command per categories. Typically all commands are composed of 2 or more words to avoid some ambiguities caused when using just one word to express a command. When using just one word, tests showed that a word could be recognised when not expected or desired because of sound from home appliances such as a TV, from conversation or from general noise. By recognising commands with more than two words we add the concept of context, and therefore although the word for an action can be recognised, a set of words must be heard before accepting the command, minimising the interference from the external noise.

Since every command can be accepted in a great number of possible ways, each command type has a similar structure which includes all the possible phrases that can trigger an action. It was difficult to tell what the action to be triggered was, based only on the possible phrases, so a field was included that specified the action that would be triggered when a phrase was detected. However, since the speech detected by the speech recognition engine made assumptions of recognised words and evaluated them with a confidence level, it was also important to capture that level. The confidence level was used for triggering actions, since only phrases with high confidence levels were accepted. To provide feedback to the user, for example when a phrase hadn’t achieved a confidence level high enough to be considered valid, this was transmitted to the user through the Graphical Interface, thereby allowing him/her to repeat the phrase and explaining what was being heard by the system. Figure 4.2(a) gives a overview of the structure that was chosen for the speech commands. There are 3 possible command types which address different contexts during speech, each of

\textsuperscript{1}Grammar, in this context, is composed of a set of words or phrases that the speech engine detects. It is used to increase system accuracy since it ignores all the words that aren’t included in the grammar, looking only for the words included in it.
which aim to cope with different situations, Fig. 4.2(b). The existing command types are Action, Navigation and Composite and, as the names indicate, the Action command type deals with device action, Navigation deals with the navigation commands going through the house hierarchy and the Composite commands are a mix of the other two command types.

**Action Commands** are composed of a set of words that express an action over a device. These commands can only happen when the user is in the lower level of the house hierarchy, which is inside a device. Commands like these depend entirely on the device operations. Each device can have different operations and can act in a different way, so therefore the grammar used adapts to these conditions. Phrases such as “Turn Off” or “Down please” are some examples of these commands.

**Navigation Commands** are used to travel within the hierarchical architecture of the house, which was also used in the graphical user interface, being used to select the devices in the house. Typically they are used to choose the place where the user wants to act over the devices, examples of which are “Enter Living Room”, “Go to Television”, etc.

**Composite Commands** are a mix of the Action and Navigation commands that were found during the Speech Grammar Tests. These commands are common in speech, making reference in the same sentence both to the desired device and to the action that will be performed on it. The actions triggered with these commands may vary depending on the context. The Composite commands can happen in all contexts and situations as opposed to the other command types which can only happen in specific situations. This characteristic increases the level of complexity in the recognition of the desired actions because when a command like this is triggered the context must be analyzed and a suitable action has to be found, for example when we are viewing the devices in a division we can say “Turn Off Television” which is a command embedded with two actions, first select the television and then turn it on. If we, instead, are viewing the available operations of the television and we say the same thing the meaning of the command is embedded with only one action, turning the TV off. These commands can only happen in the last two levels of the house hierarchy, the division and
4.3.2. Gestures Module

The Gesture module is a part of the system that deals with directional inputs. It is responsible for the capture and classification of the user’s movements regardless of the input device used. Although only a subset of input devices were used, any input device that moves the Windows operating system cursor can be used by the system without any modifications in the program configurations. This feature was intentional and is one of the most important of this system because it allows users to choose the most appropriate input device for themselves, taking into account their physical or mental constraints, without requiring changes in the system. Adaptability is highlighted in this module, which is an essential feature, especially when dealing with tetraplegic users. These users, though some may have the same level of spinal cord injury, have motor-skills which usually vary from person to person, creating a very heterogeneous segment of the population. Taking into account the characteristics of the tetraplegic population, an Abstraction Layer was implemented to ease and allow the addition of other input devices, according to each user’s needs.

**Abstraction Layer.** A huge amount of information can be extracted from a single movement, and therefore, to simplify the system, the analysis of the movement was embedded either in the input module or in the system, which then retrieves a vector that represents the direction of the user’s movement. This vector, to avoid direct communication between application and device and to provide a wider support to other devices, is expressed in cursor movement. Thus, any device that moves the Windows cursor can be easily added to the system without any modification. This choice increased the versatility of the system because we can freely test any input device without any constraint, making it possible to carry out further input testing on the system.

4.3.2.1. Input Method Recognition Algorithms

*The analysis of a movement isn’t the recognition of its many features, it is instead the recognition of the set of simple reasons that guide the user’s movements.*

This phrase was the main assumption in this module and the aim it expresses is clearly the development of ways to capture movements, focusing on the most important characteristics of these movements, disregarding others that may produce much more information but don’t clearly point to the user’s intentions.

Recognising a user’s intentions through their movements is not an easy task due to their small, imprecise and inconstant nature. However, our analysis showed that a direction may be extracted from the movement. To recognise these movements, there were two algorithms which were used: the Average Algorithm and the Discrete Algorithm.

**Average Algorithm** explored the movement based on the assumptions that movements may start with a stroke in a direction different from the user intention but the user will then make adjustments in order to correct
the direction, that will compensate for the strokes done before. This algorithm captures the slope made by each user submovement, typically captured at specific time intervals e.g. 100ms, and stores that data. When the movement ends, all the data collected will then be used to calculate the average slope which will point to the correct direction, equation (4.1). Further to the finding of the direction, the vector found will then be classified according to the available direction classes, e.g. up, down, left, right.

\[ \text{AverageSlope} = \frac{\sum \text{slope}}{\#\text{submovement}} \]  

**Discrete Algorithm** is based on the same principles as the Average Algorithm, but approaches the problem in a different way. Since the user compensates for the first stroke, which might be inaccurate, needing to do so with a set of submovements that are directed at the correct position, an approach that simply uses that first and last position of the movement may, with more precision, detect the correct direction, while having at the same time simpler calculations. The algorithm will then use those two points to create the direction vector which will then be classified according to the available direction classes.

Both algorithms use the same base algorithm which calculates the angle between two vectors from F.J. Hill [23]. An adaptation to this algorithm was made, since we needed the angle of the movement and we had just 2 points, the initial and the final point. Since considering both points as vectors was a wrong assumption, we calculated a vector using both points. With only one vector it wasn’t possible to find an angle so a standard vector was used (1, 0). The angle between those vectors was then extracted using the formula (4.2).

\[ \text{Angle} = \text{Acos(Normalize(v1).DotProduct(Normalize(v2)))} \]  

The dot product \( \langle A \cdot B \rangle \) is equal to \( |A||B| \cos(\theta) \), being \( |A| \cos(\theta) \) the scalar projection of A onto B, Figure 4.3(a). Thus, given two vectors, the angle between them can be found by rearranging the above formula to (4.2). Since \( \text{Arccos} \) had an output in the interval of \( 0 \leq \theta \leq \Pi \), which meant that it didn’t give us information about the direction of the movement (i.e. if it was ascending or descending), we used the regular slope formula (4.3) which told us, if it was positive, that the movement was ascending and if negative, descending.

\[ \text{Slope} = \frac{\Delta Y}{\Delta X} \]

Beyond these mathematical steps the algorithm used two angles. These angles expressed the inclination limits, containing areas used for movement classification. \( \alpha \) expressed the angle for the vertical and horizontal movements and \( \beta \) expressed the angle for the diagonal movements, Figure 4.3(b).

Using suitable \( \alpha \) and \( \beta \) angles, which were both 30°, and the slope formula, we were able to detect the direction of the movements in a precise and fast way.

### 4.3.2.2. Laser Module

All the input devices used, except the laser pointer, are automatically used by the operating system to do spatial-search tasks and therefore they move the operating system cursor. Because of that, we didn’t need
to develop any software for the input devices. However, since the laser pointer didn’t have that advantage, a specific piece of software, developed in C++ .Net, emulated that same behaviour.

The Laser pointer was the device that incorporated the Slash interaction mode in a more natural way. It is a fast device that requires only residual strength to be operated, and therefore it reflects the user’s movements instantly. However, in order to detect the laser dot it was necessary to disregard the parts of the image that didn’t contain the laser dot. Since the laser dot is a dot with high luminescence which is highlighted in the images sent by the Webcam, to recognise this dot in an automatic way we still had to process the image in a way so as to only show the laser dot. To achieve this purpose we used some image processing techniques based on filters. To help us in this image processing task the OpenCV library\(^2\) was used due to the variety of the material that we could use and because it was free for commercial and research use. Below, in order of their application, we explain the filters used and their objectives, Figure 4.4.

The **Gray Scale** filter was used to reduce the computation complexity of the program. Since the Webcam retrieves a frame which is a raw image with a huge amount of pixels, each of them with colour information, a huge amount of mathematical calculations must be done to process a single frame pixel-by-pixel, which consequently reduces performance. To minimise this issue, the colour information of each frame was first transformed into an 8 level grayscale which drastically decreased the information present in the frame, but at the same time preserved the laser dot due to its high luminosity, Fig. 4.4(a).

The **Binarize** filter used the grayscale frame to parse the image. It received a limit which was used to make two partitions in the image, whereby all parts of the image that had a colour higher than the limit were made white and all those below the limit were made black. This created a black and white image which, if the limit was high enough, showed the parts of the image with high luminescence, Fig. 4.4(b).

The **Erode** filter received the partitioned frame with many white spots due to ambient reflections or light sources and then applied the filter which cleaned the noise in the frame. This filter analysed the pixels around the white spots and, if the spot was small enough, turned it black. Essentially this filter cleaned all the noise

---

\(^2\)Available at http://OpenCV.sourceforge.net

![Diagram](https://via.placeholder.com/150)

(a) Visual explanation of how to find the angle between two vectors using a scalar projection. (b) Gesture recognition algorithm based on the movement angles.

Figure 4.3: Gesture capture algorithm.
present in the image, erasing the reflections and showing the place of the laser dot, Fig. 4.4(c).

(a) Grayscale filter to decrease image complexity (b) Grayscale and Binarize filter which show a partitioned black and white image. (c) Grayscale, Binarize and Erode filters which cleaned the image from noise leaving only the high luminosity laser dot.

Figure 4.4: Filters applied to an image from the webcam.

Further to the application of these filters, we used a function of the OpenCV called cvFindContours which detected the objects present in the frame. The area of each of the objects found was calculated, and then only the object with the biggest area was used. Then its centroid was discovered, telling the position of the laser dot in Webcam coordinates.

When the laser dot position was found, the process used the same algorithms for recognising gestures as for the other input devices, with the difference that when it found a direction it moved the cursor in that direction automatically and left the cursor in that position for 200ms before repositioning the cursor to the middle of the screen.

4.3.3. Modality Integration

One of the basic objectives of this system was adaptability, which meant that it should adapt to the user. An adaptable system was viewed as a system that could provide freedom of choice in input devices, allowing and delegating to the user the responsibility of choosing the right set of input devices for each occasion. Thus Gestures and Speech could be used at the same time or separated. Gestures can use a rather large set of input devices due to their Abstraction Level but Speech can only rely on sound capturing devices.

Although there is a wide range of input devices that can be used, the multimodality aspect was only tested with two modalities, Gestures and Speech, because they complement each other; if one fails the other modality can be used as a substitute. One of the first Multimodal systems was Bolton’s “Put That There” which used speech to give context and gestures to point out locations. There are already some systems that have supported this theory, such as MUSIIC [28] which scanned the surrounding ambience to create a knowledge-based speech and gesture system to control a robotic arm.

In the field of Assistive Technologies (AT), a number of AT-specific research projects have addressed multimodal input. Trevianus et al. [56] investigated the integration of speech with traditional single switch
scanning and showed an increase in selection rate, despite the limited number of vocalisations available from the subjects. Cairns et al. [8] developed a computer access environment that integrated speech, gesture, eye-gaze and pointing. To enhance computer access, Smith et al. [51] developed a system that combined speech with head-pointing to replace traditional inputs, mouse and keyboard. This system used speech for context to the text and gestures for pointing tasks. Although simple in approach, it demonstrates, like studies before it, that modalities can be mixed and the better part of each modality can be used in a system.

4.4. Information Structure

The system had to had a highly structured flow of information to deal with distinct data from various subsystems at the same time. To avoid cross-references and to simplify its complexity, a set of layers were defined, each containing a specific area of expertise which could only be directly accessed from the layers below and above. By creating these frontiers, we were able to divide the functionality between the layers, creating an easier path to solving problems, or add more functionalities without affecting the surrounding layers. Although it may seem like a standard thing to do, it drastically reduced the development of each of the systems stages, from a domotic application to a fully functional domotic system.

The basic layers, Figure 4.5, were: Physical Layer, which dealt with the details of the domotic protocols; Domain Layer, which, from a configuration file, created all the objects in the system; Application Layer, which was the centrepiece of the system, responsible for dealing with the users and for validating and solving system actions or problems; the Presentation Layer was responsible for presenting to the user what the system was sensing from its inputs and for showing the home appliances’ state.

4.4.1. Physical Layer

This layer is responsible for the connection between the system and device communication protocols. It provides an interface that is used to issue commands. It is a structured layer that provides a level of abstraction which all layers can use, protecting them from dealing with protocol specific details. It has a general interface that all domotic protocols must implement, which was done to avoid some protocol specific configurations.
These protocols are grouped into a **Protocol Manager**, which deals with the information from the protocols. In this layer a implementation of the **X10 protocol** was built, which was needed for the development of the domotic network. The protocol has some specific limitations which have already been discussed, but its most salient feature was the time taken between the issuing of a command and its effect in the actuator, which took about 1 second. Since the protocol returned an acknowledgment signal when a command succeeded, until this signal was detected the program waited as a way to synchronize the action with the interface. This choice was taken because it could give a level of assurance to the user that the command had been accepted and due to its relatively short timespan, although it still didn’t assure that the device had been actuated. The access to this protocol was made only through two arguments, Address and Function, which eased the integration in the system.

**X10 Domotic Protocol**

The X10 devices use the **zero crossing**\(^3\) of the electric current waves to send binary information. This information is sent by creating bursts of 120Hz in that no voltage zone which are then detected by other devices as bits, Figure 4.6. From then on there are a set of binary codes specified in the documentation of the protocol that are used to control the X10 devices. To act on a specific device or on the devices that have a specific house code, first an address command is triggered that identifies the devices that will be actuated and then follows a function command which will be listened to by every device activated by the last command and performed the specified action.

\[\begin{array}{cccccc}
A = 0110 & B = 1110 & C = 0010 & D = 1010 & E = 0001 & F = 1001 \\
I = 1111 & J = 1111 & K = 0011 & H = 1101 & L = 1011 & P = 1100 \\
N = 1000 & & & & & \\
\end{array}\]

(a) Start Code that precedes every command.  
(b) Address command being sent.

Figure 4.6: Transmission of data through electricity using X10 protocol.

Some discrepancies between the real protocol and the actual one implemented by the actuators were found. They were the lack of **ACK** sent when the commands were **Bright** or **Dim**, although the action itself was done, and the non-existence of some commands like **All Lights Off**. To avoid those problems we had to add some special measures to ensure the system behaved correctly.

---

\(^3\)In alternating current, the zero crossing is the instantaneous point at which there is no voltage present. In a sine wave or other simple waveform, this normally occurs twice during each cycle.
4.4.2. Domain Layer

The Domain Layer was the part of the system responsible for the maintenance and creation of the house state. It mapped the house characteristics into system objects using a XML configuration file which can be configured by the user. This configuration file was also used for configuration of some characteristics of visual characteristics of the house.

This is a complex layer, in terms of space, due to the large number of objects and functions included in it. However, there are a few basic concepts that made this system possible, which are the House and the Device objects, Figure 4.7. The House object was used to determine the house architecture, in which there were links for the Floors, Division and finally the existing devices in each division. It was extensively shown in the Presentation layer to show to the user what the current state of his/her house was. The Device object allowed the configuration of each device and its connection to the domotic network by specifying its characteristics, which were its Type, which included the commands that it could respond to, its Domotic Protocol, which specified the type of domotic technology that it was connected to, and its Address in the domotic network.

![Figure 4.7: Structure of the most important objects in the system.](image)

All the fields of these house-related objects can be specified simply by editing an XML configuration file, simplifying user personalisation and easing house changes. User can personalise not only the location-related fields but also fields related to the presentation of each device. The process of adapting the system to a new house could now be left for the user or their caretaker, decreasing maintenance costs and elevating the personalisation of the system.

4.4.3. Application Layer

Although all layers are essential for the system to succeed, they have weak links between them, increasing extensibility by separating functionality. The Application Layer, Figure 4.8, is different from all others in that aspect. In it there is most of the system’s functionality with strong links with the Presentation Layer, the Domain Layer and the Input Devices; this layer corresponds therefore to the Core Operations module described in the System Overview, see Section 4.1.

Every stimulus from any input is recorded and analysed inside the Input Analysis module which implemented the recognition algorithms, returning an action. This action will then be forwarded to the Action Validation submodule which will analyse the action, taking into account the available options in the graphical
user interface inside the Presentation layer. This validation doesn't take into account the current state of the devices because the X10 protocol doesn't have the appropriate means to assure us of the current state of the device. It can only assure us of the reception of a command. In case of a positive response from the Validation submodule, the action will be transmitted to the Domain Layer which will change the state of the objects and, if necessary, will send a command to the domotic network.

### 4.4.4. Presentation and Feedback Layer

This Layer is what the user sees from the system, it represents the frontier in which the system shows its state and from where the user bases the assumptions for his/her actions. It is composed of a Graphical User Interface and a Speech Emulating Module.
The GUI, Figure 4.9(a), was designed to simplify users’ interactions while automatically providing information about the state of the devices. It has a few components, Fig.4.9(b), each with a different function, which are the Matrix, Feedback and Status components.

The **Matrix** component is the centrepiece of all interaction. It is used to go through the house hierarchy and issue commands to the devices. Like with Gestures, here also a 3x3 matrix was the base of this design. The centre cell, unlike all other cells, is used to specify the current level of the house hierarchy by showing the appropriate image and text. All other cells map objects of the current level of the hierarchy, having attributed to them a specific movement and text which is delivered to the Application Layer to assist its Action Validation submodule. When a cell is mapped with a device, not only is the device picture seen in that cell but also its current state, by showing a picture of the device in that state; for example if a lamp is on when we are viewing the devices in a division we also see a picture of the lamp active. This way of feedback was incorporated to reduce user distraction to other components in the interface. With the matrix the user doesn’t need to look at other components, and can instead focus his/her attention on just one of them.

The **Feedback** component has the function of providing the user with feedback on his/her movements and speech. It writes the class of the movement detected and the words detected, showing the confidence level. This provides the user with a mechanism that he/she can use to perfect their interaction with the system.

The **Status** component isn’t a constant component because it changes form according to the level of the house hierarchy; for example if we are in a division, we will see the current state of the devices included in it, or if instead we are in the house level looking at all the divisions, we will see the amount of active devices present in each division. Whenever a user shows intention, this component adapts itself to that choice, always providing an extra way of feedback.

The **Configuration** has many parts used during the development stage for the developers. The first part, Fig. 4.11 1, was used for the testing of new interaction modes and to give an overview of some of the input devices that were connected in the system. The second part was used to select the metrics that would be captured in the testing phase which uses the field 3 to store the files. The fourth parts was used to start each test session and to activate the system. For simulation purposes the fifth part allowed the disconnection and connection of the system’s physical layer, thereby disabling or enabling the domotic network. This option was included because it was not feasible to install the whole domotic network in each user’s house.

### 4.5. Interaction with the System

An interaction occurs when two or more objects have an effect upon one another. In this case the objects are the user and the system, so it occurs when a user performs some action that affects the system and vice-versa.

The interaction with this system was made through some input devices and by some feedback mechanisms which required some preparation to achieve consistent results. Below we will explain the required steps to install and operate the system.
4.5.1. Setup and Preparations

Due to some principles, specified in the early stages of development, the amount of preparations was reduced to a minimum. The preparations required are related only to the installation of the domotic network, input devices and their mapping in the system.

**Installation of the X10 devices**

There are two types of X10 devices, the Computer interface (CM11) and the Actuators, both of which have a simple means of installation. The CM11 module is simply connected to the electric power of the house using the plug in the module and connected to the computer using a USB cable. All the actuators have a similar installation method, but, since they don’t need to be connected to a computer, they are only connected to the powerline.

**Installation of the Input devices**

The input devices used have some similarities, but their installation is different. The touch pad is a device already embedded in the laptops so therefore it doesn’t need any installation. The gestures used a regular 635nm wavelength laser pointer and a webcam for movement recognition. The laser pointer is a device that doesn’t need any link to a computer or other system, and it can be adapted to the user by using a special glove or an adhesive. The webcam, on the other hand, needs a link to the computer and because of that it uses a USB cable, requiring the previous installation of the proper driver. Speech also needs the installation of a microphone and, although there are already wireless sound devices in this study, a wired version was used which didn’t need any previous installation, only requiring the insertion of the cable in the correct port of the computer.

**Configuration**

All inputs of the system, except the laser pointer, are automatically detected and included in the system. The laser pointer is an exception because it requires a specific image processing program to capture its movements and therefore if the user wants to use the laser pointer he/she must run the application which, due to its cursor emulating characteristics, is automatically integrated in the system.

The part of the system that requires the most configurations is the specification of the house hierarchy and its devices. It is specified in a XML file which contains all the house-related configurations, Figure 4.10.

There is only one possible House in the configuration file. It has many optional fields, but the only fields required to achieve a correct configuration are: the existing divisions in each floor and the existing devices in each division. Each device must be configured in the Device List topic, and the most important fields are the address, name and device type. In the Device and Device Type, all fields must be completed, especially the protocol and address fields which are used in the domotic network configuration. Device Type was the object created to deal with the common properties and functions of some devices, e.g. a bed lamp has similar functions to other lamps. Each Device has therefore a Device Type that specifies its functions.

There is also a configuration necessary for the system to work, which can be configured in the configura-
Although there were 5 parts in the Configuration Panel, only the parts 3 and 4 must be specified by the user. The third part is used to identify the user and the input device used for post-analysis purposes and the fourth part was used to give a choice to the user of when to deactivate or activate the system.

4.5.2. Usage and Navigation

The usage and navigation of the system has some steps which are: Trigger of an Action, Recognition, Validation, Application and Feedback. These may seem many steps but most of them occur at the same time and just one step depends on the user, which is Trigger an Action. The user bases their choices on the Matrix group of the graphical interface, making the correspondence between the cell and the movement direction, which can be guessed by an analysis of where the location of the center cell is and the cell that the user desires to select, this being a very natural way to make a correspondence between gestures and the interface. Although the user can make movements in all directions, if the cell which corresponds to that direction is empty then no action occurs, with an error being registered in the log file, Figure 4.12.

This concept, although simple, is better explained using an example which demonstrates the changes in the interface when an action occurs, Figure A2.1. The user’s movements are directed according to his/her intentions and the mapping of the option in the Matrix. It therefore assists his/her movements. Note that when the user goes deeper in the house hierarchy the right panel, Status Panel, changes its state too, providing feedback on the device’s state in the next level of the hierarchy or in the same.

Speech uses a very simple grammar which simplifies its interaction. A user just has to say a name which is present in a cell of the matrix plus some preposition and the system will automatically select the desired cell. The most used are "Go to", "Enter in", "Turn", etc. if the recognition of these words is difficult, the user
Figure 4.12: In this example the red arrows express a directions that has no effects due to the unexistence of objects in the correspondent cells. The green arrows correspond to directions that are acceptable.

can always rephrase with a “Please” at the end or beginning of the sentence. This one extra word allows the recognition engine to achieve a greater level of confidence because it ignores other noises.
Evaluation

Through User Analysis a set of assumptions were developed which aimed to cope with some of the users’ limitations. These assumptions were assessed by means of test sessions which determined their validity. A system like this involves several technological areas united by a set of common goals, specified in 3.2, with every technology in the system tested to ensure system quality through a set of test sessions. Some experiments led to the invalidation of the proposed solutions while others led to their strengthening.

Although this is a research project, we, since the beginning, aimed to provide a platform which the users could easily use in their day-to-day life. Therefore, besides the validation of its control mechanisms, it was also necessary to validate the system as a whole for us to assess its overall usefulness. These tests aimed to discover the answers to the following questions:

Can the users control the system accurately? To control the system a set of devices were used, all using the same Matrix metaphor. Tests have already been done on some input devices to detect and avoid their major issues, see 3.4, but these tests weren’t enough because of each user’s specific characteristics. The only way to assess all the details regarding user interaction was through tests and careful analysis, which this chapter specifies.

Is the system performance enough to meet the users’ needs? Tetraplegic users have particular needs, which are similar to able-bodied individuals, but they present challenging limitations that make interaction with the unprepared existing mechanisms difficult. To meet their needs, several factors were evaluated during the task analysis phase, which culminated in the creation of a set of requirements that translated the user’s needs into system language. In principle, if the system requirements are met, the users should be satisfied with the system, but this must be assessed by user tests.

During the development of this system several evaluation phases occurred which determined the next iteration of the prototypes. In this chapter, we present not only the final evaluation tests but also some others that may explain some of the choices taken. It tries to detail the reasons why some choices persisted while other perished.
5.1. Common Evaluation Aspects

Since many testing phases were done, it was necessary to compare results that enabled the assessment between testing phases. To make that possible, a set of aspects stayed constant over each testing phase. Indeed, only the system changed in each testing phase. The parts that changed were related to user interaction and interface.

This section explains the methodology of each test section which made test comparison and analysis easier. Below is described the whole process, from the hardware apparatus to the measures captured.

5.1.1. Apparatus

A 1.7 GHz Centrino with a 15” screen laptop was used running Microsoft Windows XP Professional with a resolution of 1400x1050. The software was developed in C# .Net 2.0 and used the Microsoft Speech SDK 5.1 for speech simulation and recognition. A Labtec Stereo 242, the laptop touchpad and a TV were used for testing. For detection of the Laser dot a Creative Live! Cam Vista IM using developed C++ software was used.

5.1.2. Procedure

The tests are done in one session, with each test done individually. The session has a TV turned on at about 30-40% of its volume. It starts with an explanation of the test session followed by a questionnaire. When the questionnaire is finished, there is about 10 minutes in which the user gets familiar with the program. Then one task is explained and any of the user’s doubts clarified. The session finishes when the user completes all the test tasks with all the modalities, followed each time by a modality assessment questionnaire, and a researcher questionnaire.

1. Test environment preparations
2. Introduction and Motivation
3. Consent Form
4. Pre-Questionnaire
5. Test session
6. Acknowledgments

The task order and modality order will vary from user to user, preventing some habituation problems. In each modality or bi-modality test the first tasks presented are the even numbered tasks and when those are completed then the odd tasks are presented. When that modality is tested, the next modality tasks will be presented in inverse order to that of the last modality test, in this case first the odd and then the even.

The tasks performed are:

- T1 - Turn On the Kitchen Television (3 commands minimum)
- T2 - Turn Off the Kitchen Television (3 commands minimum)
• **T3** - Diminish the Light intensity in the bedroom twice (4 commands minimum)
• **T4** - Turn Off the lights in the bedroom and turn on the TV in the Kitchen (8 commands minimum)
• **T5** - Turn on all appliances in the bedroom (10 commands minimum)

The (multi-)modalities used are:

1. Gestures using Touchpad
2. Gestures using Laser
3. Speech
4. Gestures using Touchpad + Speech
5. Gestures using Laser + Speech

With this approach we hope that the effect of habituation will be severely reduced or at least minimized.

Depending on the users' physical constraints, if necessary the session can be divided into 2 subsessions, the first testing each input individually and the second testing the composite inputs. The test session mustn't require a high workload, and if the researcher detects a high level of fatigue he should stop the task and ask for the scheduling of another test session.

### 5.1.3. Design

All users had the microphone attached to their heads during the tests and the touchpad in a 30 cm radius from their bodies. The Camera should point to a wall or piece of paper beside the PC screen at a distance of about 1 meter from the user. The Laser pointer should be activated when the test begins and set on the users arm with a rubber band or a tape, Figure 5.1.3. Every task must start from the lower level of the house hierarchy, i.e. Divisions level, with the researcher being responsible for that change. During each task the researcher must have visual contact with the PC screen and clearly hear the speech commands without interfering with user interaction; ideally he/she must be one or two meters back from the PC screen.

### 5.1.4. Measures

Measurement is the estimation of the magnitude of some attributes of an object, in this case the user's interaction with the program. There are many types of measures, captured automatically or manually, each without the user's knowledge because it could direct the user's mood and therefore affect their performance with the program. The measures automatically captured by the program are:

- **Time** is the most common measure in almost every study. It captures the time frame or duration, in which the test was performed. When processed many statistical measures can be assessed, mainly used to evaluate performance.
- **Invalid Selections** is a derived measure which is calculated by analysis of the option selected, if the option has no content then occurs a invalid selection. This measure can only happen when using gestures since speech recognition has already a valid grammar dependent from the context.
- **Command Control Type** is the measure that explains of which modality used came the command. Since the system is multimodal the input types must be divided easing future data analysis.
The manual measures (captured by the researchers) are:

- **User Error** this measure occurs when the user makes a mistake and triggers an invalid command. An invalid command can be a gesture that the user performs and selects an empty button or, when using speech, a set of words that can't be performed in the current context, e.g. when in the house divisions level, a user says "Turn on the lamp".

- **Program Errors** occur when the user issues a valid command but the program recognises it as an invalid command or as some other command; it's the opposite of User Error.

- **Helps** occur when the user asks questions to the researchers about the interaction with the program. The users are instructed to only ask questions when it's vital to their success in finishing the tasks and therefore any questions should be about problems or doubts regarding the program interaction.

- **Command Retry** is a measure that fills the gap left by the User and Program Errors because they only capture when some action is performed in a wrong way, not capturing when a command isn't recognized. This measure occurs when a voice/gesture command has no effect on the system behaviour, i.e. not issuing any command.

When an error occurred, there were two factors that the researcher had to look for: the type of error and how the user recovered from the error. These two factors had great importance in the analysis stages because they showed the robustness and recovery of the system while capturing the reasons for the appearance of the errors.
5.1.5. Tetraplegic Participants

Although, for most of the development time, only one tetraplegic (PF) was constant throughout the development time, others joined in the latter stages. All users have regular upper body strength and coordination with the exception of the hands and arms, in which they have a small degree of control and accuracy. They have a high degree of control in the Deltoid muscle and with it they can direct the rest of the arm and hand to a desired position. Below the waist they can’t perform any voluntary movement. These users have basic computer knowledge and are familiar with Microsoft Windows Interfaces and are able to use several devices or their own hands to make selections in the interface, Table 5.1.

5.1.6. Questionnaires

During the test session questionnaires were completed. Generally, after each modality test researchers and participants filled out a questionnaire. These questionnaires tried to assess user satisfaction, the overall difficulty of each modality tested and system adequacy, see Annexes A2.3, A2.4 and A2.5.

5.2. Preliminary System Evaluation

These evaluation tests occurred during system development and were responsible for assessing some basic features of the system. The features tested were the inclusion of multimodality, the construction of a suitable speech grammar and the use of the Continuous Algorithm for classifying gestures.

Multimodality was introduced to test the effect on performance and accuracy when many modalities were used. Initially we expected similar results when comparing with unimodality. However, we also expected some confusion when using some or many devices at almost the same time.

The Continuous Algorithm was chosen at this stage because we suspected that a slope analysis alone throughout the movement wouldn’t be able to correctly classify the movement. Tetraplegic users are widely recognised for their lack of accuracy in bodily movements However, we suspected that some parts of the movement were more important than others and therefore we used this algorithm to provide us with such data.

<table>
<thead>
<tr>
<th>User ID</th>
<th>Age</th>
<th>Injury</th>
<th>Injury Type</th>
<th>Cause</th>
<th>Qualifications</th>
<th>Control Mechanisms</th>
</tr>
</thead>
<tbody>
<tr>
<td>PF</td>
<td>31</td>
<td>C5</td>
<td>Complete</td>
<td>Dive</td>
<td>12th grade</td>
<td>Arm Stick</td>
</tr>
<tr>
<td>NC</td>
<td>34</td>
<td>C5</td>
<td>Complete</td>
<td>Fall</td>
<td>University Degree</td>
<td>Keyboard Buttons</td>
</tr>
<tr>
<td>HD</td>
<td>41</td>
<td>C6</td>
<td>Complete</td>
<td>Accident</td>
<td>12th grade</td>
<td>Hand/Arm Motions</td>
</tr>
<tr>
<td>SA</td>
<td>26</td>
<td>C4</td>
<td>Complete</td>
<td>Accident</td>
<td>University Degree</td>
<td>Eye Tracker</td>
</tr>
</tbody>
</table>

Table 5.1: Users technological profile table.
Sound-based modalities are well-known for having problems when dealing with noisy environments. However, these problems are enhanced when the grammar used for speech recognition is large. By reducing the available grammar, we hoped to increase its accuracy and performance, but we also needed to know how much we could reduce the grammar. By capturing the words emitted by the user to trigger an action, we could create a proper grammar, suitable for the users.

5.2.1. Participants

5 participants (4 male 1 female) of average age 25.2 participated, one being tetraplegic (PF) and 4 able-bodied. All participants were familiarised with WIMP GUI’s and the most regular input device used by them was the mouse for the able-bodied, and touchpad for the tetraplegic. All users used a standard keyboard to text entry tasks.

5.2.2. Procedure and Design

The procedure and design taken in this experiment were similar to that which were explained in section 5.1.2 5.1.3. However, the Laser wasn’t used because it showed some challenges regarding its interaction and therefore we used this test section to validate other choices regarding the algorithms. User PF leaned on the bed with the laptop by his side, while other users did the test while sat in a wheelchair with the laptop in front of them. In this test session just the touchpad and the microphone were used.

With this experiment we expected a minimum number of commands, see (5.1), issued to complete the tasks. A number of commands above that value should point to user or program errors.

\[
\text{MinimumCommandsCollected} = 28 \times \text{MinimumTaskCommands} \times 3 \times \text{ControlModes} \times 5 \times \text{Users} = 420
\]  

(5.1)

5.2.3. Results

Tests showed some problems or difficulties in controlling the system, which were expressed by high error rates, Figure 5.2(a). The speech was clearly the most efficient in terms of the number of operations and error rates (4%) followed by a much larger sum for gestures (24%) and their conjunction (28%) for the able-bodied. The tetraplegic population achieved a large error rate in gestures (74%), with their results for speech and speech mixed with gesture modalities having a close relationship to those for the able-bodied population.

The gesture discrepancy between able-bodied and tetraplegics can be explained by the physical constraints that the tetraplegics have; able-bodied people can achieve a finer-grained and more constant movement, so creating less errors.

The able-bodied population’s worst performance was when using both gestures and speech together, with an average time of 27.1 seconds per task, but there wasn’t much difference between all the modalities used. The better results achieved were when using only gestures, with an average time of 19.8 seconds, which can be explained by the regular use of the touchpad by these users. With tetraplegics, gestures showed a
serious lack of performance, with a 78.16% decrease when compared to able-bodied people, which can be explained by its 73.5% error rate. Speech performed well and had similar results to those for the able-bodied population, but when speech and gestures were mixed the high error rate for gestures decreased its overall performance by 22.14% for the able-bodied population and 42.24% for the tetraplegic, Figure 5.3. The times clearly expressed the difficulty characterised by the error rates.

The Touchpad was the first to be tested, which may have influenced the results because of user inexperience regarding the program, Figure 5.2(b). Users made many errors while using only Gestures due to recognition problems of the algorithm, a consequence of the highly responsive movement recognition which issued a command whenever the cursor changed position, leading to unwanted and unexpected commands. Another problem was the inability to change the cursor position when the cursor was on the screen border, which had a larger effect since users couldn’t make their desired movement.

The user assessment questionnaire and the researcher questionnaire, both done after each user test session, revealed that the physical and mental effort required to operate the devices was low to non-existent. All users evaluated Gestures and Speech from comfortable to very comfortable and also as comfortable for Speech and Gestures mixed. The researcher questionnaires showed a lack of adequacy in the gestures.
5.2.4. Discussion

Results showed high user satisfaction and easy interaction, although further studies are needed to increase the confidence level of the statistics. For the target user population the system has shown great value, though requiring some adjustments. The Gestures need improvements such as the automatic repositioning of the cursor after each command and an adjustment to its sensitivity to reduce its responsiveness. Voice and Gestures can be used in an environment when voice alone fails to achieve acceptable results, an environment that wasn’t achieved in the tests, although some noise was simulated. Speech seemed to be the most viable modality for both user groups at this early stage of development, although gestures for able-bodied people have a slightly better performance. The huge amount of errors from gestures disabled any attempt at using the Continuous Gesture Recognition Algorithm. Therefore future user tests should test the Discrete Gesture Recognition Algorithm and compare both algorithms.

5.3. Performance and Accuracy Evaluation

Taking into account the results from the last test sessions, some modifications were made regarding the interaction modes. The high user rate in gestures was unexpected. However, refinements were made in the Discrete Algorithm using the Slash interaction mode.

This test session evaluated the interaction of the users with the system, comparing the advantages and disadvantages of each modality.

5.3.1. Participants

In this test session we focused on the tetraplegic user group and therefore they were the participants used in this test session. Four tetraplegics with an average age of 33 years old (SD=6.27), two with C5 injuries, one with C4 and the other with C6 injuries. All users were familiar with the common Windows interfaces which they used regularly through a set of mechanisms, see Section 5.1.5.

5.3.2. Procedure and Design

This experiment tested the whole system, exploring all its options and their combinations. The design and procedure of this experiment was detailed in the sections 5.1.25.1.3. The user PF did the tests in bed leaning to the right, unlike the other users who did them sitting in a wheelchair due to physical limitations.

With this experiment we expected a minimum number of commands, equation (5.2), issued to complete
the tasks. A number of commands above that value should point to user or program errors.

\[
Minimum\text{CommandsCollected} = 28\text{Commands} \times 5\text{ControlModes} \times 4\text{TetraplegicsUsers} = 560
\]  

(5.2)

5.3.3. Results

The results from these tests are divided into 3 parts: Accuracy, Performance and User Adequacy/Usability. These three factors provide a clear basis that allows a correct overview of the system’s potentialities and limitations. Although they are linked, by analysing each of them separately we can produce objective conclusions that could benefit the system as a whole. Some problems occurred during the experiment; the user PF experienced severe difficulties in controlling the laser pointer and therefore the tests using that device were excluded with this user, though they remained with the other participants.

5.3.3.1. Accuracy

Capturing the accuracy of the system was a joint effort between the researchers and the system using the analysis manual and automatic measures. Although some errors were easy to detect, such as the selection of a empty cell, others, such as the incorrect calling of a object, had increased difficulty. There are two distinct analyses taken from this study, an analysis of the modalities and inputs, Interaction Analysis, and an analysis of the relation between the number of commands of each task and the errors of the users, Command/Task Analysis.

Interaction Analysis

The interaction through the touchpad clearly showed an average error rate of about 0.67% which was clearly lower than that of all other devices. On the other hand, the laser was the device that performed the worst with 9.52%. As expected, speech showed some problems with a 6.47% error rate. However, by mixing speech with the touchpad we achieved a 13.79% lower error rate than when compared with only speech. When mixing speech with the laser pointer we were able to reduce the error rate by 18.75% when compared with only the laser pointer, see Figure 5.4.

Dividing the errors by its origins, the tests showed that all speech-related modalities had a similar number of user errors, which varied from 4.25-4.75. Indeed all the gesture-related inputs showed that the number of program errors was greater than user errors, while modalities that included speech showed the opposite, user errors were more numerous than program errors. Gestures through the touchpad had a better accuracy with only 0.22% of user errors and 0.45% of program errors. Speech showed an average of 4.25 user errors and 3 program errors per task, but when mixed with the touchpad, there was a decrease of 50% in program errors at the same time as a increase of 11.76% in user errors. When using Speech and Laser, results appeared to be promising because there was a drastic reduction in errors when compared to only the Laser; results showed a reduction of 38.09% in program errors but a increase of 18.18% in user errors.
In terms of helps, some surprising results were found. Speech and Touchpad had the most help requests with 3 helps per task, while Speech and Laser had the least requests with only 1 request per test.

By linear regression we were able to find a tendency line which could point us to the values that could possibly be achieved if other compositions of modalities were included. The number of helps, as expected, decreases when modalities are mixed which reflects the increase of the understanding of the system and its interaction modes by the user.

**Command/Task Analysis**

The number of minimum commands to trigger correctly the task increases as the number of the task rises. However, each task was created to test particular aspects of the interaction. To make a contextualized analysis of the tests, we must understand the most important features of each task. Tasks from 1 to 3 explored the directions inside an appliance. Task 4 explored interactions in different divisions which increased the navigation commands. Task 5 had the greatest number of operations and explored many diagonal movements inside one house division. Tasks 1 to 3 are similar, with the difference of a higher use of the diagonal movements in Task 3 and therefore results should be similar. Task 3 showed a 280% increase in user errors with 3 errors when comparing to Task 1 and 2. Task 4 had the most program errors with 4.2 but the number of helps was low with 1.6 helps on average. Users in Task 5 asked for help more than in all the other tests, which resulted in similar user and program errors, see Fig. 5.5.

**5.3.3.2. Performance**

The users’ performance while interacting with the system was captured by the completion times of each task. We expected similar times for all users but their particularities created some heterogeneity in the results, see Fig. 5.6. Results varied from 28.82 to 44.34 seconds on average with the laser pointer being the input with
the worst results and the touchpad with the best results.

The most constant user, NC, had times from 35-45 seconds in every modality. All others, surprisingly, achieved better times when using the Touchpad, which is a gesture-based device. Speech also demonstrated the results shown by the error analysis. Its times were reduced by 6.29% when using speech with the touchpad and by 5.05% when using speech with the laser pointer. When comparing the times taken by the laser pointer in combination with speech, astonishing results appeared, with a reduction of 15.24%. These results clearly showed a reduction in times when comparing use of mixed with use of only one device or modality, except in the case of the touchpad, which had great results and in fact the combination with speech increased its times by 28.71%.

One user surprised, achieving better results with the laser than with the touchpad. Indeed all users that used the laser pointer, except HD, showed better results using this device than with speech from 19% to 21%. User HD, despite his poor results when using only the laser pointer, achieved better results in all others, which clearly pointed to his better motor-skills due to his lower spinal cord injury.

### 5.4. Usability and User Satisfaction

By means of the questionnaires we could assess the users’ satisfaction with the inputs used and with the system. All users pointed out the laser pointer as the least comfortable input device, classifying it as intermediate in the comfort aspect. Speech was found to be the easiest and the most comfortable to operate. The mental effort to operate the system was described as residual or not-existent in all devices or combinations. The physical effort was only an issue when using the laser pointer, which, due to its wide movements, was
found to require medium effort.

Researcher’s questionnaires revealed that speech and speech-related modalities had several problems because of the use of incorrect English and some problems appeared due to the speech feedback which was captured again by the system, creating a loop which was difficult to solve.

5.5. Overall Discussion

From the analysis done, the touchpad clearly showed the best results both in terms of performance, Table 5.2, and in error rate with a superiority of at least 35% when compared to all the other devices and combinations. During the tests, the tasks made using gestures that navigated more in the house hierarchy and the tasks which used more diagonal movements presented more errors than other tasks with a larger number of commands. When asked after the tests, users were both impressed and satisfied with the performance of this device. Speech presented medium results, demonstrating high error rates and medium performance. However, it was the preferred choice of the users, and its high error rates might be explained by noise and speech feedback from the system. The speech feedback, in some situations, was captured again by the microphone creating a loop in which the command explained by the speech feedback was received by the system as a new command.

Although some problems existed when using speech, its use in combination with the laser demonstrated a decrease of 5.05% in time when compared with only speech and 38.09% in program errors while a increase of 18.18% in user errors when compared with the laser. When mixing speech with the touchpad, results showed a 6.30% decrease in time and 50% in program errors while having a increase of 11.76% in user errors, compared with speech. Clearly, Speech showed that when mixed with other gesture-based devices, it could improve their accuracy and performance. The Laser pointer had the worst results whether in performance, in
accuracy or in user satisfaction, though user NC achieved even better results with it than with the touchpad, in terms of performance. This fact expresses the heterogeneity of this user group.

By linear regression we found the decreasing tendencies of helps, as the tasks have higher commands less helps are needed. These results, although for it to be expressive a larger user group is needed, showed that the users with some training were getting used to the system. The level of performance which could be achieved by the users should be similar to the performance achieved by the user PF, which clearly shows in speech and in touchpad better performances when comparing to other users, because he was the user that most closely followed the system development.

**Comparison between Test Stages**

This comparison has the objective of making the comparison between two distinct evaluation phases. Since only one user was constant in both phases, namely PF, the comparison will be centred on this user to evaluate not only the effect of familiarity with the program but also the changes that occurred. The biggest difference between these stages was the gesture recognition algorithms, Continuous in the preliminary evaluation and Discrete in the final evaluation.

In terms of error rates, both in speech-based and gesture-based input, there was a drastic reduction, Table 5.5, from 50.61% to 99.09%. In the only gesture-based input, the touchpad, these results can only be explained by the use of the Discrete Algorithm instead of the Continuous and by the increased familiarisation with the control mechanisms. Speech also had high error reductions, which is clearly a consequence of the customisation of the user-centred grammar, which adapted the speech commands to the user’s vocabulary.

<table>
<thead>
<tr>
<th></th>
<th>Preliminary Error Rate</th>
<th>Actual Error Rate</th>
<th>Comparative decrease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Touchpad</td>
<td>73.5%</td>
<td>0.67%</td>
<td>99.09%</td>
</tr>
<tr>
<td>Speech</td>
<td>13.1%</td>
<td>6.47%</td>
<td>50.61%</td>
</tr>
<tr>
<td>Speech + Touchpad</td>
<td>35.6%</td>
<td>5.78%</td>
<td>83.76%</td>
</tr>
</tbody>
</table>

Table 5.3: Error Rate Comparison between the Preliminary Evaluation and the Final Evaluation.

The reductions in times, Table 5.5, were also expressive, especially with the touchpad, which achieved results 73.03% lower. The dramatic decrease in errors rates was reflected in performance, enabling the users to achieve better times. Although the decreases were significant, the small difference between the Touchpad
<table>
<thead>
<tr>
<th></th>
<th>Preliminary Times (sec)</th>
<th>Actual Times (sec)</th>
<th>Comparative decrease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Touchpad</td>
<td>90.7</td>
<td>24.43</td>
<td>73.07%</td>
</tr>
<tr>
<td>Speech</td>
<td>32.2</td>
<td>25.89</td>
<td>19.60%</td>
</tr>
<tr>
<td>Speech + Touchpad</td>
<td>54.8</td>
<td>42.66</td>
<td>22.15%</td>
</tr>
</tbody>
</table>

Table 5.4: Time Comparison between the Preliminary Evaluation and the Final Evaluation.

and the Speech of 5.6% in performance, allied with the higher comfort and physical effort of Speech, could be the reason why users pointed to Speech as the preferred input.
Conclusion

We have presented a system that enables tetraplegics to control their household appliances. Through User Analysis, a set of objectives were created, taking into account the users’ needs and the environment. It was found that the users needed a system easily adaptable to a changing environment while providing the control mechanisms suitable to deal with each user’s limitations, using low-cost technology.

The functionality of the system used the X10 domotic protocol through a set of actuators to control the home appliances. Although some time occurred - about 1 or 2 seconds - before the actuators performed the actions, this wasn’t an issue for the users. The limited functionality provided by this protocol seemed to be enough to cope with most of the tasks performed.

The real challenges of this dissertation were centred on capturing the user’s intentions and translating them into correct system actions. The interaction therefore had to cope with many challenges from both system and each individual user, which led to the development of a multimodal interface. Gestures and Speech were chosen due to their suitability with Tetraplegics.

To cope with our objectives, some control approaches were developed, all based on a 3x3 matrix, using the developed Discrete algorithm for gesture recognition and an adjusted grammar for Speech, which boosted the performance and accuracy of the system. Results showed high user satisfaction and easy interaction, although studies with more users are needed to increase the confidence level of these statistics. For the target user population the system has shown great value and has the potentiality of increasing user independence, presenting times for activating 4 devices in a bedroom of under 1 minute for all users. Gestures using the touchpad presented the best overall results, while the laser pointer presented the worst results, a fact that could be explained by its free-hand interaction. Tests also showed that diagonal gestures had more errors when compared to all other gestures, which could be explained by users’ limited motor-skills. The overall system interaction with these control interfaces was considered simple, requiring only residual physical and mental effort to be operated.

The system successfully coped with the objectives initially specified. However, it wasn’t possible to conclude that it would be advantageous for Tetraplegics due to a reduced test sample in terms of users and due to the non-existence of long term studies that could prove the overall usefulness of the system in a tetraplegic’s day-to-day life. Since users were pleased with the characteristics of the system, its being installed in a
tetraplegic's house, in the hope of increasing their autonomy and providing enough data to support the main objective of this work was achieved, which was to "return to tetraplegics the ownership of their houses".

6.1. Future Work

In future iterations, some system limitations should be tackled in order to extend the potential user range and to study/enhance its usefulness. Below there is a description of the aspects that should be explored in future:

**Permanent Use Scenario**

Intensive use was an important scenario that wasn't captured by the test sessions. These scenarios usually hide other problems due to their uncontrolled circumstances, not present in usual test sessions. After the installation in a tetraplegic's house, or ideally in many houses, for an extended period of time, the data gathered should provide us with enough information to assess the overall usefulness of the system, highlighting its limitations.

**Portability**

The system suffers from lack of portability because it's based on a laptop. Although the control Interfaces used - mainly speech - can provide us with some level of portability, the laptop still needs to be close to users, limiting its action range.

By providing a mobile extension that can be connected to the existing system we could bring the system closer to users, allowing them to control the system using a small device. A device similar to a PDA would be ideal since in it there is already embedded a gesture control interface and the necessary mechanisms to use voice.

**Control Interfaces**

In terms of control interfaces, a few were tested by the tetraplegic participants. However, other modalities and inputs must be studied to ensure the suitability of the control mechanisms for these users.

The interfaces that could fit these users more properly and that should be studied in future studies are the EMG and BCI. These two interfaces have a wide potential user range and were developed taking into account the needs of the physically challenged, so therefore by using them we could extend the potential user range of the system.
Bibliography


A1.1. Task Analysis

For making the task-analysis the "11 Questions" approach was used [15], it is a standard task analysis method that aims to the construction of a User-Centered Design. Below there are the most relevant questions to the definition of the system.

**Who are the users?** The users are tetraplegics with low spinal cord injuries (C5-C8). These users have some kind of control over their trunks and arms, but have no control over their hands. They live in home with one or more caregivers present most of the time, exceptionally (one day a week) there are some periods of time that they are leaved alone, in the afternoon or saturday night.

These users have daily contact with the Windows GUIs, interfaces that they operate daily, therefore they are considered expert users.

**What tasks do they perform actually?**

To operate the house appliances all the tasks done require caregiver assistance, they haven’t got enough dexterity and mobility to reach the devices. To operate with personal computers 33% of the users can interact with the computer using the mouse, although with difficulty, while 66% of the users use the touchpad of a laptop computer. All users use the keyboard for text-entry operations.

They can receive calls, start calls and write a SMS, although very slowly, with a mobile phone.

**Which are the preferable tasks?** All users demonstrated interest in controlling their surrounding appliances remotely. The most important devices were: TV, lights, blinds, heating system and sound system.

**Where are the task performed?**

The tasks are performed inside their homes which are adapted to their mobility needs. The users spend most of their time in bed, 8-12 daily hours, having the need of changing position in every 2-4 hours. Occasionally users are in an electric wheelchair (8 hours maximum) 2-3 times a week.

**What's the relationship between user and information?**

They use the laptop to access information due to its portability, since it can adapt to user bed positions.
They use a regular keyboard and touchpads. To contact other persons or be contacted they use a mobile phone.

**What devices do they have?**
The users have, in front of their beds a large TV and within a 3 meters radius a laptop. The house has the regular home appliances: Lights, TV, Sound System.

**What is the frequency of each task?**
The most frequent tasks, 3 or more times a day, are operations with the TV, Sound System, mobile phone, lighting and computer. With the exceptions of the mobile phone and computer, all other devices are operated via caregiver.

**What are the time restrictions?**
The users have restrictions imposed by the physical disability. In the bed or in the wheelchair users have to change body position in every 2-4 hours.

**What happens if anything goes wrong?**
If anything goes wrong the caregiver/s will come in aid of the user and do the tasks for them.

### A1.2. Initial Evaluation

The performance evaluation of a pointing device has some difficulties due to inconstant, oscillating performance and accuracy over time. There are many factors that can influence the results, a large percentage of them the participant doesn’t even notice. Emotional and Physical tiredness seem to dictate the performance of each individual during the tests. These factors make the collection and analysis tricky at best.

Beside those particularities there are common features between classes of devices and its control methods and that’s what this paper specify, using a set of measures.

### A1.2.1. Measures

For specifying a movement or set of movements several metrics were used, some of them considered standard [25] and other proposed by other researchers [33].

**Movement Time** \((MT)\) is the most traditional performance which is the "time that users spend moving the pointing device" [52]. According to Fitts’ Law, \(MT\) is predicted with the following formulation:

\[
MT = a + b \times ID
\]

Where \(a\) and \(b\) are device constants discovered experimentally by linear regression, being the intercept the slope respectively.

There are two types of \(ID\) (index of difficulty) \((A1.2)\), the nominal \((ID)\) and the effective \((IDe)\). The Shannon formulation was used for the ID calculation.

\[
ID = \log_2 \left( \frac{D}{W} + 1 \right)
\]
\[ IDe = \log_2 \left( \frac{D}{We} + 1 \right) \tag{A1.3} \]

\[ We = 4.133 \times \sigma \tag{A1.4} \]

The effective index of difficulty (A1.3) uses the movement end-points to perform an adjustment for accuracy using the standard deviation (\(\sigma\)) to create the effective target width (We) (A1.4) [52].

**Target Entries (TE)**

Ideally, one target entry per target is the perfect score because it expresses an accurate control of the pointing device. This measure can assess over the accuracy of movement, specifying the amplitude as excessive or accurate.

**Task Axis Crossing (TAC)**

In every test there is a starting shape and a finish shape, the centers of those shapes are used to create a line that is the Task Axis. This measure, proposed by Mackenzie et al.[33], captures the number of Task Axis Crossings to characterize the trajectory the participant took when approaching the finish target.

**Horizontal Direction Shift (HDS)**

If the pointer changes trajectory in the x-axis, ex. moves left and then right or vice-versa, a HDS occurs.

**Vertical Direction Shift (VDS)**

This measure is similar to HDS with the difference that the axis used is now the y-axis, it then records the vertical oscillations.

The Horizontal and Vertical Direction Shift in the perfect movement will be zero, which tells us that the movement was linear. Those measures will help to characterize participant tremor in the movement path.

**Movement Error (ME)**

Originally proposed by Mackenzie et al.[52], this measure calculates the average absolute deviation from the task axis. It gives an overview of how far from the task axis the movement was performed. The formulation is:

\[ ME = \frac{\sum |y_i|}{n} \tag{A1.5} \]

**Offset Error (OE)**

Similar to ME, OE is a regular average distance between the pointer from the task axis. By analysis of this measure, one can assess if the user has a tendency of making a movement over or bellow the task axis.

**End-point Distance to Target Center (EDTC)**

This measure specifies the distance between the target center and the selection point. These measures can be used to infer user laziness, in easy tests, or precision.

For this experiment there were two times recorded AT and ST.

**Acquisition Time (AT)**

The Acquisition Time is the time taken to acquire the target, i.e., the time taken until the pointer first enters the active shape.
Selection Time (ST)

This measure is the time taken from the beginning of the movement until the selection of the active shape.

The Acquisition Time and Selection Time ideally will have a correlation between them since one precedes the other there will be a $\Delta t$ corresponding to the time taken to do a selection “Click”, about 100ms for the mouse. If this correlation doesn’t exist one can assess that there isn’t a balance between Precision (ST) and Performance (AT) times.

A1.2.2. Results

Below there is the explanation of some of the remaining metrics capture during these test sessions.

A1.2.2.1. Errors

There were 11 errors during the tests, due to user habituation to target disposition, users tended to move the cursor before the next active shape was activated. Other error situations happened in the 4 bit difficulty test using the Laser Pointer in Absolute mode, 5/11, one person had severe difficulties and didn’t was able to finish the first test. The overall error rate was 53/1122 which is 4.7%.

A1.2.2.2. Path Analysis

The total end-point distance was similar for all devices, except Laser Mouse Joystick which had the higher end-point distance of all devices (23 pixels). Keyboard uses mostly vertical and horizontal lines having a trajectory distant from the Task Axis mainly in diagonal targets. Trackball had trajectories mostly bellow the Task Axis while lasers had higher variability. The device that had trajectories closer to the Task Axis was the touchpad followed closely by the mouse, Figure A1.1.

Every device entered the targets, Figure A1.2, in average 2 times except Laser, as expected. Laser Mouse Absolute (LMA) had an average of 9.7 (SD=15.997) and Laser Mouse Joystick (LMJ) had 5.56 (16.831) which represents a 43% decrease although still distant from mouse with 1.97 (3.074).

Laser Mouse Absolute was the device that crossed the Task Axis the most, 3 times more than the trackball
which had a mean of 12.9 (69.686).
Vertical and Horizontal Direction Shifts were relatively similar between devices.

Figure A1.2: Path variations per device.

A1.2.2.3. Performance Comparison
LMJ has shown the best results comparing to the LMA, although both have much slower times than all other input devices. Mouse and other pointing devices showed low times which could be explained by their indirect interaction, Figure A1.3.

A1.2.2.4. Subjective Results
The ISO 9241-9 Assessment Test results demonstrate that lasers are not comfortable and difficult to use under those conditions, analysing the fatigue indicators, medium-high, one may infer that the lack of comfort could be influenced by physical fatigue. LMJ reveals that it is more uncomfortable than LMA which could be a result of being the last device being tested in every session.

A1.2.2.5. Time Comparison between Laser Acquisition and Selection Times
To make a contextualized assessment of the laser pointer we also needed to know what was the difference between selection and target acquisition times and if they stayed constant in the tests. No data was found to prove that the selection times were made in a constant period of time after the target acquisition.

By linear regression the Fitt's constants $a$ and $b$ of A1.1 were found and are displayed in Figure A1.4.

All the comparisons between the times of both devices are statistically significant, Figure A1.5 has the results from paired t-tests.
Figure A1.3: Mean Selection Time, Target Acquisition Time of the devices.

Figure A1.4: MT $a$ and $b$ determined experimentally per device.

Figure A1.5: Laser Times compared statistically.
A2.1. EasyHouse: Test Script

I am David Graça, Msc student from Lisbon Technical Institute, specializing in assistive technologies development. My work aims at the development of a system that eases the access to home appliances for persons with motor difficulties, using gestures and/or speech. To improve system reliability and performance I need your support in a set of simple tests and questions. All your personal information will be treated with confidentiality and won’t be divulged.

During the tests, only if necessary you should question the researchers, they will have a duration of about 40 minutes and shouldn’t require much workload or user tiredness. The data provided by these tests will make possible a better user characteristics knowledge and adequacy, therefore approaching my work to the user.

A2.2. Consent Form

During these tests some attachments may be required such as a small rubber band to control the Laser pointer or a microphone. You will be asked to perform a set of tasks by the researcher through many input devices. Keep in mind that at anytime you can refuse to continue the test session. The system and the will collect data regarding your interaction such as times and commands issued. The data processing will be done anonymously disregarding any personal data, it will be treated with a random user identifier. Please read the statements bellow and sign the ones that you agree with. Thank for your participation in this study.

1. I agree to perform the evaluation session described and that the results are logged in text files.

Print Name:

Signature:

Date:
2. I agree that the session is videotaped to further analysis by the researchers.

Print Name:

Signature:

Date:

3. I agree that the multimedia (photos and video) captured during the evaluation are used by the researchers in dissemination events and publications (thesis, articles, press).

Print Name:

Signature:

Date:

A2.3. Pre-Questionnaire

Personal Data

1. Name:
2. Contact:
3. Age:
4. Damaged Vertebrae:
5. Complete or Incomplete Vertebrae:

Motor Abilities Classify the following table from 0 to 2 the type of control that you can achieve in each part of the body:

<table>
<thead>
<tr>
<th></th>
<th>Central</th>
<th>Left</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td>Face</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eyes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neck</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shoulder</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forearm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hand</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trunk</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finger</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**User Experience**

1. Have you ever used a Mouse to navigate in a computer?
2. Have you ever used a Touchpad to navigate in a computer?
3. Have you ever used a Laser to navigate in a computer?
4. Have you ever used a Speech recognition program to navigate in a computer or with other purpose?
5. Have you ever used an application that controls your home appliances?

**A2.4. Questionnaire after each subsession**

Classify from 1 which is none or very low to 5 representing very good or very high the following questions:

1. The physical effort required for operation was:
2. The mental effort required for operation was:
3. The response time was:
4. General comfort:
5. Overall, the input device was:

**A2.5. Researcher Assessment Questionnaire**

Classify from 1 which is none or very low to 5 representing very good or very high the following questions:

1. The physical effort required for operation was:
2. The mental effort required for operation was:
3. The adequacy level of the test regarding to the test subject was:

**A2.6. User Interaction Example**

Example of a regular user interaction in the EasyHouse user graphical interface, Figure A2.1.
Figure A2.1: Example of interaction. The user will turn on the lamp in the bedroom. The green arrows were here used for better understanding of the movements done to activate an option.