Universidade Nova de Lisboa
Faculdade de Ciências e Tecnologia
Departamento de Informática

Dissertação de Mestrado em Engenharia Informática
2º Semestre, 2011/2012

Gesture Based Persuasive Interfaces for Public Ambient Displays

Ricardo Manuel da Silva Salvador Nº 32514

Orientadora
Prof. Doutora Teresa Romão

24 de Setembro de 2012

Gesture Based Persuasive Interfaces for Public Ambient Displays

Nº do aluno: 32514
Nome: Ricardo Manuel da Silva Salvador

Título da dissertação:
Gesture Based Persuasive Interfaces for Public Ambient Displays

Palavras-Chave:

- Sistemas de computação persuasiva
- Mudança comportamental
- Public Ambient Displays
- Sensores 3D
- Interfaces Gestuais

Keywords:

- Persuasive Systems
- Behaviour change
- Public Ambient Displays
- 3D Sensores
- Gesture Interfaces

Acknowledgments

I'd first like to express my appreciation to my advisor Dra.Teresa Romão. Thank you for giving me the opportunity to work in the DEAP project, and of course for your time, patience and guidance.

I'd like to thank all of the members of the DEAP project, particularly Dr. António Eduardo Dias and Pedro Centeiro, for their constant and priceless feedback throughout the dissertation. I'd also like to thank the CITI- IMG team for sharing their resources and their stimulating work environment. I give great thanks to Bárbara Teixeira for her amazing contribution on the graphic design of the MAID. I also have to thank my brother António Salvador for is last minute proof-reading of this document.

I am also very grateful to my colleagues and friends that provided moral and technical support throughout this journey: Vítor Pereira Rui Felizardo, João Araújo, Filipe Araújo, Rui Rosendo , among others.

Last but not least, I'd like to thank my family, specially my parents, for their ever present support throughout my life. And of course a special thanks to my girlfriend, Marta Sousa, who gave me her unconditional support during this process.

This work was funded by Fundação para a Ciência e Tecnologia (FCT/MEC), Portugal, in the scope of project DEAP (PTDC/AAC-AMB/104834/2008) and by CITI/DI/FCT/UNL (PEst-OE/EEI/UI0527/2011).

Resumo

Esta dissertação estuda como Public Ambient Displays (PAD) podem ser usados como ferramentas para atingir uma mudança comportamental, através de tecnologia persuasiva.

Para atingir os objectivos desta dissertação, um sistema PAD interativo, chamado Motion-based Ambient Interactive Display (MAID), foi desenvolvido. MAID tem como objectivo motivar mudanças comportamentais em relação ao consumo de energia doméstico, através de um interface de jogo persuasivo baseado em tecnologia de reconhecimento de gestos. O protótipo desenvolvido guia o jogador através das várias divisões de uma casa, onde este terá que descobrir o que está errado de forma a praticar as acções correctas para poupar energia, usando gestos semelhantes aos que usaria no mundo real para desempenhar a mesma acção. O sistema retorna informação de contexto sobre as repercussões de cada tarefa desempenhada, para que os utilizadores se apercebam das consequências das suas acções

A implementação do MAID é baseada numa estrutura altamente configurável e modular construída para o efeito. Permite que administrador afine a aplicação de acordo com as necessidades do local de instalação, através de ajustes às propriedades do ecrã, modificação dos conteúdos de imagem ou até alterando o próprio guião do jogo. O sistema de guião de jogo é flexível o suficiente para permitir que esta estrutura seja reutilizada para outros estudos, para além do âmbito desta dissertação

O MAID foi sujeito a testes de utilização, para demonstrar que é possível criar um sistema de PAD persuasivo, usando métodos de interacção seamless, com a tecnologia actualmente disponível e usá-lo para sensibilizar os utilizadores para a causa, levando a mudanças comportamentais.

Abstract

This Master thesis studies how Public Ambient Displays (PAD) can be used as a tool to achieve behaviour change, through persuasive technology.

In order to reach the goals of the thesis, an interactive public ambient display system called Motion-based Ambient Interactive Display (MAID) was developed. MAID is driven to motivate behaviour changes regarding domestic energy consumption, through a persuasive game interface based on gesture recognition technology. The developed prototype guides players through the different rooms of a house, where they have to find out what is wrong and practice the correct actions to save energy, using similar gestures to the ones they would use in real life to achieve the same goals. The system provides feedback regarding the consequences of each action, in order to make users aware of the consequences of their actions.

The implementation of MAID is based on a purpose built, highly configurable and modular framework. It allows the administrator to fine tune and tweak the application to the necessities of the setup location constraints, by adjusting basic display properties, change image content or even modify the scripted gameplay itself. The scripted game system is flexible enough to allow the repurposing of the framework, beyond the previously defined theme, for future studies.

The MAID was subjected to user testing, in order to show that it is possible to create a persuasive PAD interface, using seamless interaction methods, with the currently available technology, and use it to spread awareness of a cause, leading to behaviour change.

Table of Contents

List of Figures

List of Charts

1. Introduction

The human influence over the environment has been a widely discussed and documented issue for decades, since it may cause irreversible damages. An important issues is the increase of greenhouse gas emissions, caused for example by use of fossil fuel as an energy source that are resulting in a global climate change, increasing the average world surface temperature at an alarming rate (Chart 1.1) (Global Warming Articles News and Facts, 2011). It is important to save energy, since it helps to reduce the consumption of environmental resources and the CO2 emissions, contributing to a sustainable environment.

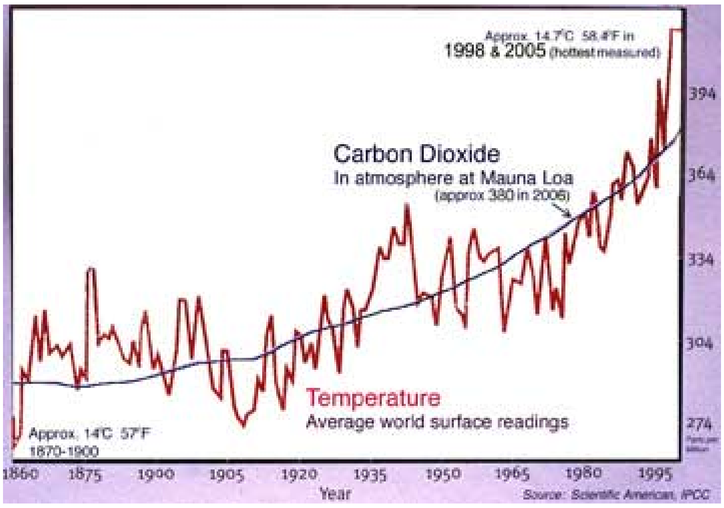


Chart 1.1 - Carbon Dioxide emissions vs. average world surface temperature

The destructive pattern of human behavior will have to change dramatically very soon, and in some cases, it may even be too late to avoid irreparable damage. This may be the case for the Antarctic ice sheets, where specialists believe “there is a 50% chance that widespread ice sheet loss "may no longer be avoided" because of greenhouse gases in the atmosphere”, which will “raise sea levels by four to six meters”. This would result in the flooding of coastal regions, endangering cities like Tokyo, London or New Your City (Adam, 2007).

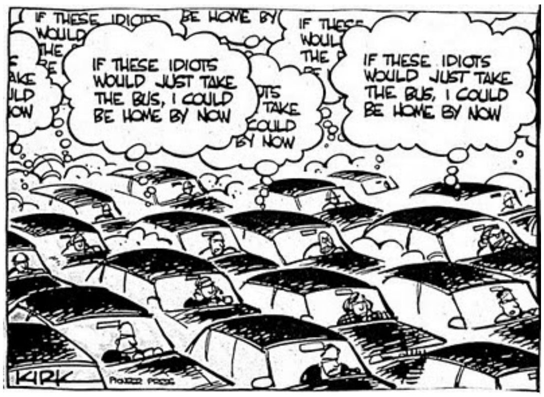


Figure 1.1 - Why do we leave it up to others?

Throughout the years, environmental concerns have been publicized in many different formats in practically every media channel available, and now-a-days almost everybody that has access to the media is informed of the current situation, but one thing is to be informed, and another is to take action. Sometimes it is hard to grasp the impact of the actions of only one person in a global cause, but why should we all expect others to take action, when we are not willing to (Figure 1.1)?

This thesis shall explore new means of delivering a persuasive experience to the users, in order to motivate a behavior change. This experience will be in the form of an interactive public ambient display, taking advantage of the latest technology available in the field of Human-Computer Interface.

1.1. A brief introduction to Human-Computer Interfaces

Human-Computer interaction is a melting pot of behavioural sciences, design, computer science and other fields (Human–computer interaction, 2011). It involves the study of how a user interacts with a computer and the evaluation of the degree of success obtained. To do so, the machine side of the interface draws wisdom from techniques in computer graphics, operating systems, and programming languages among other. As for the human side, graphic design, communication and cognitive psychology theory are the dominant fields.

Figure 1.2 - Apple Lisa Commercial

Traditional interfaces are based on a computer keyboard, mouse and screen for hardware and a window based operative system interface, popularized by Xerox and Apple (Figure 1.2) in the early 80's in their widely available machines, derived from the early light pen (Figure 1.3) and sketchpad projects (Myers, 1998). This interface paradigm turned into an industry standard, for better or worse. It made multipurpose systems possible, many kinds of applications could be controlled throw it, from text editing to assisted drawing. In the following years many new concepts for HCI have been documented, but none have been able to reach the popularity of the keyboard, mouse, display and window based OS. One of the reasons is that the use of this model "stuck" with the users, just like entering a car and expecting to find a steering wheel, a shifter and a set of pedals.

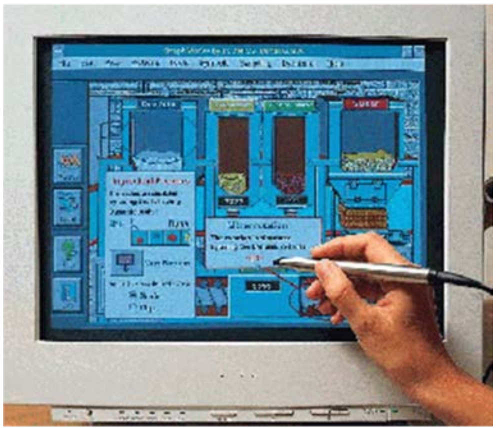


Figure 1.3 - Xerox Light-Pen

Another reason is related to the price of this technology, a mouse and keyboard setup is inexpensive nowadays and can handle most of the tasks that are needed in a common PC setup.

This scenario has changed in recent years, with purpose build interfaces, engineered to fit the needs of the target device. We can now find practical capacitive touchscreens that do not require the use of a stylus, in common mobile phones and Pads, running operative systems specifically design with that mean of interaction in mind. Nowadays, the hardware can fit in a very slim device, like the Apple I-Pad (Figure 1.4), with a battery that lasts up to 10 hours, keeping the total weight of the device between 500g and 800g. So, the interface has to take into account the fact that this device is engineered not to be physically hooked up to any peripheral devices, and capacitive display fits the need for an affordable technology that allows a user to take full advantage of the device anywhere he wishes to use it. It is now common to see people using touch-screen based devices on the street, but this kind of technology still lacks the precision needed for tasks that require typing and pointer based interaction, needed for the average office job, where the keyboard and mouse combination is still the popular option.

Figure 1.4 – Apple I-Pad tablet

Public Ambient Display interaction is one of the cases where the keyboard and mouse interface are just not practical enough. In a public environment, like a street, for brief spontaneous interaction, the current and most popular solution is the use of interactive displays, like touch-screens or even multi-touch screens, like the Microsoft Surface (Figure 1.5), persuading users to explore the system through a clean and hassle free interface technology. But there are many limitations in these technologies, like the wear and tear from the many users. Another limitation is the size the screen, limiting the maximum number of simultaneous users.

In this thesis I shall explore the HCI problem posed by the PAD Interaction and tackle it using pervasive technology. The technology that was chosen was depth sensors, to detect user gestures through 3D processing. This technology had recent developments that resulted in new hardware tools available on the market.

Figure 1.5 - Microsoft Surface

1.2.Motivation

Many environmental problems occur due to our erroneous and exaggerated use of Earth resources. Recent studies suggest that, within USA, 22% of domestic energy consumption could be eliminated, if people engage in more appropriate behaviours, such as replacing traditional light bulbs with low-consumption bulbs or adjusting thermostats and turning off lights when leaving rooms (Laitner, 2009). Each one of us can contribute to reduce domestic energy consumptions and consequently CO2 emissions. Small individual actions may seem insignificant, but together everyone's efforts can have a great impact on the environment.

Huge amounts of resources are used every year in these kinds of actions, with efforts being made by many governmental and non-governmental institutions. Most of these effort may inform their targets, but do not motivate a real and lasting behavior change. Most of the successful examples of these campaigns try to pass their message through the everyday life of their targets, with a constant and direct influence, imposing a behavior change.

These kinds of measures are applied on many causes, like anti-tabaco campaigns, with the objective of reducing the number of active smokers. These campaigns can go as far as Nazi Germany in the early 1930's, where Adolf Hitler's personal distaste for tobacco, among other reasons, fueled some of the first anti-tabaco campaigns. Nowadays, it is almost impossible for a smoker not to know the consequences of this habit, but it is common knowledge that in many cases there is no real behavior change until he is faced with the consequences. With that in mind, in the early 2000's a new simple strategy was devised, where each package of cigarettes was branded with a direct and harsh phrase or image that demonstrated the consequences of smoking, so that each time a smoker held the pack, he would be reminded of the consequences of smoking (Figure 1.6). It is hard to evaluate the results of this kind of campaign, since there are so many other campaigns and other factors that influence the results. But this format of delivering a persuasive message has been implemented in many countries throughout the world, with a variety of messages and never stopped evolving in order to increase its effect.

Figure 1.6 - Cigarette pack warning labels

Some initiatives to promote a behavior change, for the good of the environment, have also taken a more active path. One example is the new waste management policy in the city of Lisbon: Non distinct domestic trash used to be retrieved every day from the street trash


Figure 1.7 - Recycling bins

bins. In order to influence people to recycle their domestic trash, the street bins were replaced by appropriate recycling bins (Figure 1.7), and the retrieving of the trash is done by categories (Plastic, Paper, Glass and Non distinct), at different days of the week. This action had a direct influence in the life of the citizens of Lisbon, leading to a behavior change where people included recycling habits in their daily routine, making it a lasting one.

The modern day dependencies on technology, and the globalization phenomenon associated with the dawn of the internet, lead to new means for persuasion and behavior changes. Computer based technology plays an intricate role in our day to day life, from the office computer to the cash register in our local department store. These technologies are breaching new boundaries by the second, connecting people all over the world, through structured social networks that facilitate human social interaction through the internet. These modern social networks, like the well-known Hi5, Facebook or Bebo, allow the users to do more than the traditional profile with pictures and basic friendship bond. They can be used for direct chats, competitive virtual gaming, publishing of many art forms and so many other things.

It never ceases to amaze how the versatility of modern technology can be put to good use, in order to allow the delivery of messages in a fun, practical and creative way. Technology can be used to influence people's behavior, regarding the environmental concerns, in order to lead them to a sustainable usage of the resources of our planet.

1.3. Goal

The Goal of this thesis is to research new means of human-computer interface in the context of Public Ambient Displays (PAD). Exploring new ways to deliver a persuasive experience, through the development of a PAD prototype, motivated by environment related concerns. By using seamless means of interaction, the interface was designed to persuade the user to implement a behaviour change. The objective behind the persuasion methodology is to lead users to realise how a simple action can have a big impact, when applied in a large scale, and to show how it contributes to the big picture.

The PAD implementation is based on a purpose built, highly configurable framework that enables the PAD installation to adapt to its surroundings. It allows the administrator of the PAD to configure as much as possible, not only to tune the interface to a site specific set of constraints, but also to edit the content of the user interface itself. The interface technology is also very flexible, allowing the PAD to be installed in a variety of conditions: bright or dark, noisy or quiet, empty or crowded. It does not need any type of hand held controller or any contraption that the user has to wear. With these constraints in mind, the prototype goes beyond the simple novelty gadget with single event demonstration logic. It enables the study of persuasive technology in different settings, where the PAD configurations are fine-tuned for each setup at each particular location.

The prototype also underwent user studies to evaluate the system's usability, gameplay and reactions of users.

1.4.Contribution

This thesis was developed in the scope of Project DEAP (PTDC/AAC-AMB/104834/2008), which aims to introduce new paradigms for environmental awareness, helping to motivate citizens to become more environmentally responsible in their everyday lives and engaging them in environmental preservation activities.

One of the main contributions is the development of a versatile public ambient display framework, which can be used as a mean to deliver a persuasive interface, with few constraints regarding the installation's surrounding environment. It can be used as a tool to study persuasion methodologies in a public setting. It applies a new interaction methodology for PAD, using current widely available technology, asserting the viability of this kind of interfaces as a mean of effectively deliver a behaviour changing persuasive experience to the user.

Another contribution is the use of gesture based interface methodology, exploring the correlation between the simplicity of the gestures with the corresponding action in the real world. Accentuating the idea that small gestures can have a deep impact.

The system developed in the scope of this thesis was demonstrated, as a creative showcase, at the 8th International conference on Advanced in Computer Entertainment Technology (ACE), originating a publication on ACM DL (Salvador & Romão, 2011). A full paper will also be presented at the ACE 2012 and published by Springer (Salvador, Romão, & Centieiro, 2012) (see Annex 3).

1.5. Document Structure

This Master Thesis report is structured as follows: section 2 describes the state of the art related to the main areas covered by the presented work; section 3 regards the early prototype development; section 4 presents the MAID prototype; section 5 deals with the planning, realization and results of the user studies, and finally, section 6 presents conclusions and directions for future work.

2. Related Work

This chapter summarises the state of the art of the main areas covered by this thesis. Section 2.1 is dedicated to the design principles of persuasive systems and their influence on behaviour change. Section 2.2 shows how persuasive technology can be applied to environmental causes and current work in this field of study. Section 2.3 illustrates how an online social network can be used as a mean to deliver persuasive messages. Section 2.4 is dedicated to interactive public ambient displays design principles and current work in this field of study.

2.1. Persuasive Systems

”Fifteen years ago, there were relatively few examples of persuasive technologies in our lives. The web wasn’t ubiquitous and software wasn’t designed to change behaviour. But today (…) we are surrounded by digital products designed to change what we think and do” (Fogg B. J., 2009) . Persuasive technology can be found in many different forms, from shopping suggestions on sites like Amazon, which persuades users to buy items related to the ones that the user selected, to modern massive multiplayer games, that persuade us to never stop playing, through social pressure, because the success of other users in the game environment are depending upon cooperative initiatives. Or even modern cars, persuading the users to have a driving behaviour that favours low fuel consumption, by displaying many statistics of fuel consumption in the heads-up displays. Persuasive technologies can be categorized by their functional roles. B.J. Fogg proposes three “basic ways that people view or respond to computing technologies” (Fogg B. J., 2002). Computing products can take the role of social actors, tools or media (Figure 2.1):

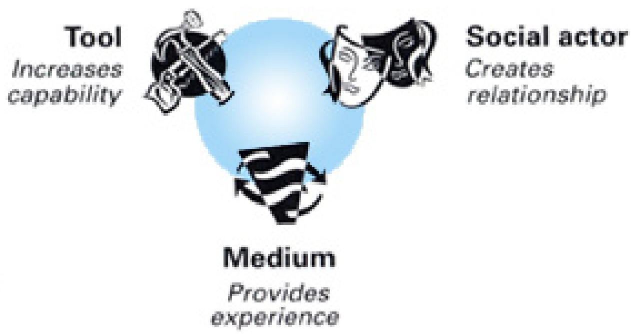


Figure 2.1 - Persuasive Technology: Roles

As a tool, persuasive technologies can be used to increase people’s ability to perform a target behaviour, by making it easier to achieve the goal or/and by organizing the tasks needed to reach a goal. For example, the ticket vending machines, located in almost every train station, use persuasive technology as a tool, in order to guide a user through the task of buying a ticket, from selecting the type of ticket and the destination of the user to collecting the money and issuing a receipt. This way the user is persuaded to avoid the ticket vending office.

As medium, persuasive technology can be used for rehearsing behaviour, in order to persuade the user to change it. For example, there are some driving schools that are equipped with educational driving simulators, with the goal of rehearsing everyday driving situations, so that the user can learn the proper behaviour for the simulated scenarios.

As social actors, an application will try to interact with the user , because “when people perceive social presence, they naturally respond in social ways—feeling empathy or anger, or following social rules such as taking turns” (Fogg B. J., 2002). For example, in many modern video games, the user is motivated to progress in the game story through interaction with virtual companions, whose virtual lives are depending on our success in the game.

Nowadays there are so many forms that persuasive technology can take that many researchers are investing on studying the dynamics and principles of persuasive technology (Fogg B. J., 2002). With that in mind, B.J. Fogg created the eight step Design process (Fogg B. J., 2002), using successful examples of persuasive technology as a basis for developing new forms of persuasion. Starting with simple behaviours as targets for persuasion, identifying current examples of successful persuasive technology in that area, implementing the persuasive system based on the examples and expanding on results obtained. It is quite easy nowadays to build a persuasive system, using tools like social network systems, but many fail because people don’t understand what factors lead to behaviour change (Fogg B. , (A), 2009) (Fogg B. , (B), 2009) (Penny & Underwood, 2008). Behaviour changes can be achieved through the conciliation of three main factors: Motivation, ability and trigger (Fogg B. , (A), 2009): The system must motivate the user to act towards a target behaviour; it must have the ability to accomplish the target behaviour; there also must be a trigger so the user can act upon it to accomplish the target behaviour (Chart 2.1).For example, I would love to learn to play piano, given the love that I have for music, I am motivated to practice and I have enough ability to practice just by following an instructional book. The trigger for me to practice playing is having every week a friend to practice with.

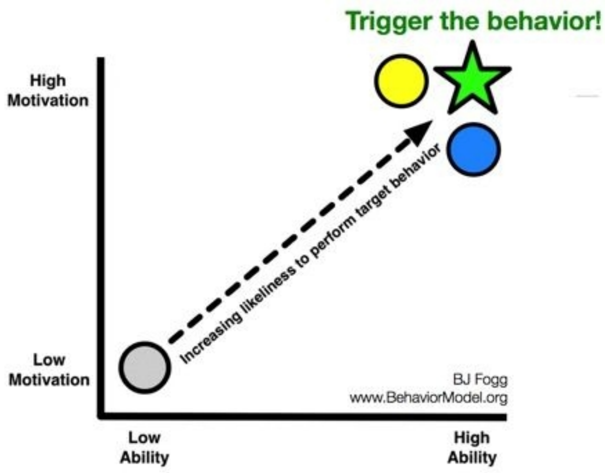


Chart 2.1 - The Fogg Behaviour Model has three factors: motivation, ability, and triggers.

These simple principles are applied as the foundation blocks for the development of the thesis prototype, in order to build a persuasive system. Using modern new means of interaction, together with a stimulating and visually appealing design to supply motivation to the user, combined with the use of natural and context appropriate gesture based interface to illustrate how simple and accessible the target behaviour is. The final component would be the supplied feedback, the impact of the data supplied by the system should act as trigger.

2.2.Persuasion methods applied to Environmental Causes

Persuasion technologies can and have been applied within several contexts, such as commerce, health care (de Oliveira, Cherubini, & Oliver, 2010), education (Revelle, Reardon, Green, Betancourt, & Kotler, 2007) and environment (Froehlich, et al., 2009) (Reitberger, Ploderer, Obermair, & Tscheligi, 2007) (IJsselsteijn, de Kort, Midden, Eggen, & van den Hoven, 2006) (Miller, Rich, & Davis, 2009) (Stein, 2008) (Sutaria & Deshmukh, 2008) (Takayama, Lehdonvirta, Shiraishi, Washio, Kimura, & Nakajima, 2009) (Spaargaren & Van Vliet, 2000).

Several approaches explore the use of games as persuasive tools to influence players to change certain behaviors, there are four roles that should be considered as a perspective to design persuasive system (Midden, Kaiser, & Teddy McCalley, 2007,):

As an intermediary, where the technology, that is used for attaining a goal, defines the ecological impact.

As an amplifier, where technology amplifies the human potential to attain goals.

As a determinant, where behaviour is shaped and activated on the basis of the affordances, constraints and cues provided by the technological environment

Figure 2.2 - Faucet with proximity sensor

As a promoter where technology is designed to influence behavioural choices.

In the effort to produce behaviour changes, there are many possibilities for the implementation of persuasive systems. One of the solutions is the purely technological approach. One example of this approach is the use of devices that limit the users' behaviour to an acceptable level, like faucets that have proximity sensors, so that the water flow is only turned on when the users have their hands in front of the sensor (Figure 2.2). This is also one of the most ineffective approaches, in terms of producing a real and lasting behaviour change, given that the user is forced to act in a certain way instead of maintaining the target behaviour from its own free will. This is mainly due to human nature, people tend to resist to changes: “resistance to new systems and negative experiences (…) has frustrated the high hopes of innovative technologies” (Midden, McCalley, Ham, & Zaalberg, 2008).

Figure 2.4 - (top) a polar bear with lots of ice (bottom) a polar bear with little ice.

Using the four roles mentioned earlier, technology can be used as an effective motivational tool to help people visualize the consequences of their behaviour and explore the effects of their actions. Persuasive systems can be planed beyond the traditional level of specific appliances or systems and aim towards integrated systems that monitor multiple resources and advise the users on how to adjust their behaviour or invest in more efficient equipment. These systems can even go beyond the scope of the common household and coordinate efforts within the society, coordinating contributions. One recent example of a persuasive system aiming towards behaviour changes was based on the users’ attachment to a virtual pet “persuading individuals to change their behaviour, and maintain those changes over time…” (Dillahunt, Becker, Mankoff, & Kraut, 2008). In this experiment, a group of individuals were selected to test an application where a social actor, a Polar Bear, was used to create an attachment. The actor’s environment was the polar ice caps, and the ice would melt, making the actor sad, or grow, making the actor happy, according to the behaviour of the user (Figure 2.4).

Figure 2.3 - Persuasive displays designed, and deployed in field studies

There are also other recent experiments comprising a more sensor based approach, with the objective of contributing to the users’ awareness of their domestic consumptions and the respective consequences, aiming towards behaviour changes (Loke, Singh, & Le, 2008). An example of this approach is the UpStream project, with the goal of “immediate reduction in water use …through unobtrusive low-cost water flow sensing and several persuasive displays” (Kuznetsov & Paulos, 2010). This project experimented with different means of passing the information to the user, regarding current water usage, in order to persuade them to reduce their water consumption. These means were separated in two groups, Ambient Displays that use sound or light patterns to alert the user, and numeric displays that show a precise value (Figure 2.3).

Figure 2.5 - The Power Agent game

The results lead to the conclusion that both methods are effective, but numeric visualization was less effective than ambient. This was probably “due to the display’s lack of a clear-cut threshold for appropriate consumption: participants did not see an appropriate stopping point”, so the users did not have a clear notion of progress and immediate goals that visual feedback would supply. These conclusions and results were taken in consideration for the development of this thesis, so that the prototypes take advantage of the available forms of delivering feedback to maintain the user informed on their status and drawing clear objectives for the interaction, using visual as well as audio output to deliver an engaging gameplay.

Regarding domestic energy consumption, some applications have been developed. Power Agent (Bang, Gustafsson, & Katzeff, 2007) is a mobile pervasive game for teenagers, designed to influence their everyday activities and electricity usage patterns in the household. Power Agent is played on standard Java-enabled mobile phones (Figure 2.5) and uses real power consumption data collected via existing metering equipment used by the energy company. The underlying design idea is to let players (one for each house) compete in teams (cooperate with their families) and learn hands-on how to save energy in their homes. Power Explorer (Gustafsson, BÃ¥ng, & Svahn, 2009) builds on the previous research on Power Agent and uses a similar technological set-up, but focuses on real-time feedback (Figure 2.6). The game design teaches the players about the consumption of their devices and encourages them to adopt good habits. Some other applications focus on providing feedback based on real-time domestic energy consumptions measurements (Jahn, 2010) (Mattern, Staake, & Weiss, 2010) (Reeves, Cummings, Scarborough, Flora, & Anderson, 2012).

Figure 2.6 - The main view of Power Explorer (the habitat).

Although these approaches seem to produce positive results, they rely on the use of mobile devices and the deployment of specific energy consumption devices (connected to the home central electric power meter or appliances), which involves additional costs and set-up. We aim at using players' free time in public spaces, such as shopping centers or waiting rooms, to both motivate and educate them towards home energy saving behaviors. Our approach is based on the assumption that a change in behavior requires both motivation to act as well as knowledge of what to do.

2.3.Persuasion Through Social Networks

In the last few years, a new form of persuasion emerged, the mass interpersonal persuasion (MIP), through the evermore popular social network systems. One of the main contributors to this phenomenon was Facebook, a social networking service, that "created a new way for third parties to develop and distribute interactive applications (web apps) …" (Fogg B. J.,(A), 2008). Applications on Facebook are available with a wide range of objectives, including sharing contents and social gaming.

Figure 2.7 - Barack Obama's Facebook Profile

The social networks have turned out to be a prolific way to be heard. Nowadays many people that seek publicity (Figure 2.7) use large social networks to achieve it. For instance, the Facebook social network has a functionality that allows the users to promote causes, that other users can join and support. This resource has been used even by non-profit organizations to promote causes like animal rights, environmental concerns, cultural movements and so many others.


Figure 2.8 - Facebook's Farmville App

The social networks systems are always adding new features to their lists. One of the popular features, that is now available in all the major social networks, is the support for third party applications. These applications can bind external systems to the power of the social graphs.

One of the means of interaction that are gathering popularity in the social networks is the competitive virtual games, constructed as third party applications. These games are simple in nature, like FarmVille, where users compete to achieve the most beautiful and productive farm (Figure 2.8). Each game has a set of achievements and levels that the users compete to achieve. Each achievement is shown in the users profile/page, like a professional athlete and his set of cups and medals in his trophy room. Users invest considerable amounts of time using these applications, to reach these achievements.

This typical competitive environment of these applications can be used as a mean of persuasion that would lead to behaviour changes, using the influence that social networks can have over users to steer them towards a more environmentally sustainable living. According to B.J. Fogg (Fogg B. J., (A), 2008), there are six components in MIP:

- **Persuasive experience** - The Persuasive experience in the MIP is characterized by social influence. Many applications on Facebook are an example of this, by action on social pressure. For example, one of the most successful applications, the "Top Friends" application, sends this invitation message to the users friends: "John Smith has added you as a Top Friend! Does John Smith make your Top Friends?". This message is designed to persuade the user to use the application by leveraging the fact that the message comes from a friend that the user probably trusts.
- **Automated Structure -** One of the most important aspects of MIP is the automated structure of the system, constantly delivering a persuasive experience. So once a persuasive experience is designed and implemented, it is applied repeatedly with no need of human assistance.
- **Social Distribution** -Persuasive experience and automated structures have been used in many occasions, like web application based marketing, but one of the main innovations in persuasive systems, regarding MIP, is the adding of social distribution where the persuasion influence can cascade through the social network, from friend to the friends of the friends and so on. One of the most important examples of this is the growth of the social network systems, with the invitation system, where a user invites his friends to join the network, which cascades to the friends of the friends and so on.
- **Rapid Cycle -**Another component is the rapid cycle, "what this means is that the time between invitation, acceptance and a subsequent invitation needs to be small" (Fogg B. J., Mass Interpersonal Persuasion: An Early View of a New Phenomenon, 2008). One of the best examples is the Facebook group system, where a group can represent a cause, and users can join the group to support the cause. These groups tend to gather huge followings trough rapid cycle spreading of the information.
- **Huge Social Graph** -MIP cannot exist without a huge Social Graph. The impact

of the components described depends on the size of the social graph that in the case of networks like Facebook, tie millions of people together. In the future, there may even be some form of connection between different social graphs, linking specific components of different social networks, expanding the social graph even further.

- **Measured Impact -** The final component would be the measuring of impact, so the user can observe the effects of their actions. One of the ways is through social proof, where the success of, for example, a group on Facebook can be confirmed by the amount of followers the group acquired. Another way is through statistic information. Using the same example, a user that joined a group can get information on the average number of users that joins the group per day. One of the most useful ways of measuring the impact of an application, for the creators, is by monitoring the user interaction with it. Counting how many times a user pressed a particular button, how much time a user spent on a window and so on. This information is priceless for the developers to adjust their product to fit the users' behaviour.

Mass interpersonal persuasion has the potential to make an individual be heard by millions, without the limitations of the normal means of mass media, decentralizing the power of persuasion. We can already see these effects through the influence that individuals have had with blogs and online videos, and that was only the starting point.

2.4. Interactive Public Ambient Displays

The technology and interaction methods for public ambient display design have gone a long way from the simple non interactive display that only showed useful information in a static way. Nowadays, PAD play an important role in many people's lives, with this technology being applied in many different environments.

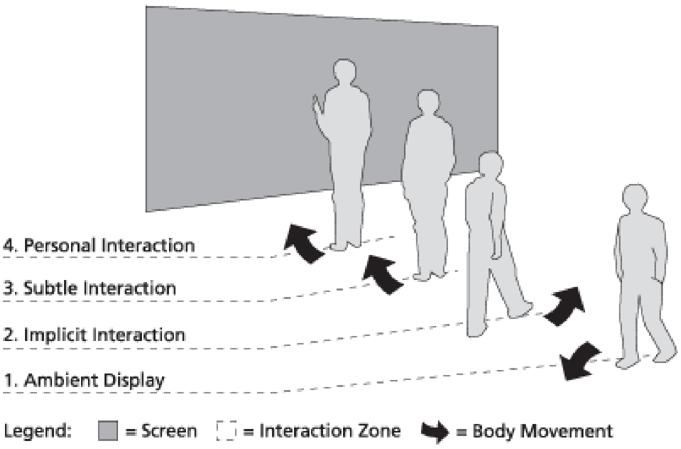


Figure 2.9 - PAD Interaction Zone

Some more simple PAD are applied in environments like airports, for displaying flight related information, where users congregate to be informed of departure time, gates and state of their flights, and a stock market PAD were investors can gather most of the data related to the market's up's and down's. Other, more complex PAD's, support human interaction, like interactive class room boards, that assist teachers on their daily lessons.

The PAD system paradigms are evolving to the point of contemplating "exchanging specific information with individuals as they pass by" (Figure 2.9) (Vogel & Balakrishnan, 2004). As the information networks begin to solve privacy issues, and users start to trust and relay upon them, visualization of personal information will be possible through PAD, avoiding the need for Laptops and other devices to access personal information. Although there is still a long way to go, many efforts have been made to breach this barrier.

Following these ideas, in order to guide the designing process of an interactive PAD system, there is a set of design principles we should follow (Vogel & Balakrishnan, 2004):

- An ambient display should be designed according to the context of its surroundings, considering the aesthetic aspect and the interface reactions, in order not to attract too much attention or to pass unnoticed.
- The information transmitted by the PAD should be comprehensive, even if the user has to interact with the system to obtain more details.


Figure 2.10 - Prototype sharable interactive ambient display.

- The display should communicate with the user in a socially acceptable way, avoiding a too invasive approach. In order to maintain the ambient nature of the display, the interaction with the user should be short, limiting the interaction to information gathering as opposed to activities.
- The system should not require prior training, aiming at immediate usage, and for more advanced interaction, the PAD should demonstrate in a straightforward way the method for interaction.
- Multiple users should also be able to use the system, either individually or simultaneously (Figure 2.10).
- The PAD should reason which personal information should be appropriate to display in a public environment, avoiding displaying information that is inadequate, like user's e-mails or personal messages.
- And finally, methods for users to control their privacy, in the PAD, should be made available, so that a user can choose what will be displayed in the system when he interacts with it.


Figure 2.11 - Shadow Reaching: An interaction technique

Based on these principles, Vogel & Balakrishnan also created a framework for interaction phases on PAD, which range from the distant implicit public interaction to the up-close explicit personal interaction, with fluid inter-phase of transitions (Figure 2.10).

The use of a keyboard and mouse is unrealistic for the modern PAD systems. Future ambient display technology should merge with its surrounding as perfectly as possible. In order to achieve that, the means of interactions should be as natural as possible. There are many on-going projects, studying new ways to interact with PAD's (Jacucci, et al., 2010) (Shoemaker, Tsukitani, Kitamura, & Booth, 2010) (Eisenhauer, Lorenz, & Zimmermann, 2008). For example, there is a project focused on Whole-Body Interactions (Shoemaker, Tsukitani, Kitamura, & Booth, 2010), where the whole human body is used for interaction, based on the projection of the users shadow on a wall, in order to detect the users body gestures (Figure 2.11). Another project studies the possibility of separating the user interface from the PAD system, in order to create a common interaction platform that would enable the developers' to concentrate on the design of the PAD system itself and not the interface method (Eisenhauer, Lorenz, & Zimmermann, 2008). This common interface should be modular enough to accept many combinations of different input devices, like cameras and pointers. The concept of separating the graphical interface from the HCI makes more sense than ever, with technological advances creating new HCI devices with different approaches, the interfaces developed for a HCI framework can be applied to new input device by only updating the HCI side.

Figure 2.12 - A participant using a space game in a public setting

According to (Hardy, Rukzio, & Davies, 2011) gestures are a viable method of interaction with public digital. Their work con-tributes to the understanding of how people perceive, respond to and interact with interactive gesture based public displays outside the controlled environment of a research lab (Figure 2.12). Their findings support both the observe-and-learn model (Rubegni, Memarovic, & Langheinrich, 2011) and the Honeypot effect (Brignull & Rogers, 2003) and we followed the recommendations that were appropriate for our case study. A public display allows for a broader dissemination of a message, since besides the user directly interacting with it, all members of the audience can receive the output from the public display.

Implementing a successful gesture-based interface also requires the challenging task of designing a set of effective gesture commands that allows users to interact in a natural, familiar and effortless manner. This is particularly important in the context of sporadic interaction with public displays when users have limited time to learn and explore the application (Vatavu, 2012).

Our gesture set was designed to resemble the corresponding real world actions, so the gestures would be easier to understand and recall.

3. Early work in the development of the Prototype Concept.

3.1. First concept designs:

This chapter will describe the first stage of the creations of the thesis prototype. At this point, the main concern was to run through the first ideas for the concept, test their viability and assert the necessary technology.

3.1.1. Approach

Given the goal of building a public ambient display system, as a mean to deliver a persuasive interface, through innovative and seamless human-computer interface technology, the first step taken was to research the available technologies that could fit the needs. This stage of development was really focused on testing technologies, in order to find the one that is the best match, given the goals established, with particular infuses in the flexibility requirements (*see chapter 1.3*). The first immerging ideas' include the exploration of computer vision techniques that allowed the users to interact with the PAD without the use of any direct physical contact with an HCI device.

3.1.2. Technologies applied

The first approach was based on computer vision, using the OpenCV library for real-time video feed processing. The image capturing was done by an inexpensive webcam.

OpenCV was chosen not only because of its list of features, but also because it is very well documented and supported, a priceless advantage in this early stage. The interface was developed using the open-framework libraries, supplying important abstractions that allowed the quick development of an early prototype.

3.1.3. Concept

The basic interface concept was to create simple games, based on environmental related topics, that would draw a non-optimized scenario that the user could explore. Persuasive technology works as a medium, so the user can learn and practise the target behaviour by optimizing the scenario (Fogg B. J., 2002).

The first game concept was based on domestic environmental issues, like excessive power consumption, wasting of water, recycling among other. The user would be presented with a domestic scenario that contained a set of situations that should be corrected, like a dripping faucet, that should be fully closed, or an incandescent light bulb that could be replace with an economic light bulb. A mixture of direct numeric and ambient information would be studied, to work out which combination had the best effect (Kuznetsov & Paulos, 2010), for example, a combination of dials that show the current power and water usage could be combined with ambient light effect in the game scenario, that with the game progress would go from a greyish game environment to a more “happy” and green lighting.

The user would also be given the possibility of publishing the results on a social network, like Facebook. Each game would have a presence on the social network, in the form of a third-party application, and each user could participate in spreading the word about the project, tapping into the established social graph. The use of the social graph would allow us to further publicise the cause, it would also increase the effectiveness of the persuasive technology. The expansion of the persuasive experience from the PAD to the social network would be subjected to study.

3.1.4. Explored Technologies

For this early concept, I decided to limit the interaction to hand gestures, since 2D visual recognition technology, using a webcam, clearly had limitations to the shape identification process, due to lighting, shading and colour contrast. So, by using hand gestures, I minimized the impact of these issues, by narrowing the field of vision, with the hands occupying most of the captured frame (*Figure 3.1*), obtaining the best correct detection rate possible.

In order to process the computer vision information, OpenCV libraries were applied to the camera data stream, to extract the hand gesture information. The extrapolation of the hand position was mainly based on the detection of the convex shape supplied by OpenCV. This shape is created by the space between the user’s fingers, combined with colour segmentation based on skin colour, it is possible to calculate the hand contour (*Figure 3.1*). With the position of the most essential points of the hand, like the finger tips, it is possible to detect gestures, based on hand/finger position and direction shifting.

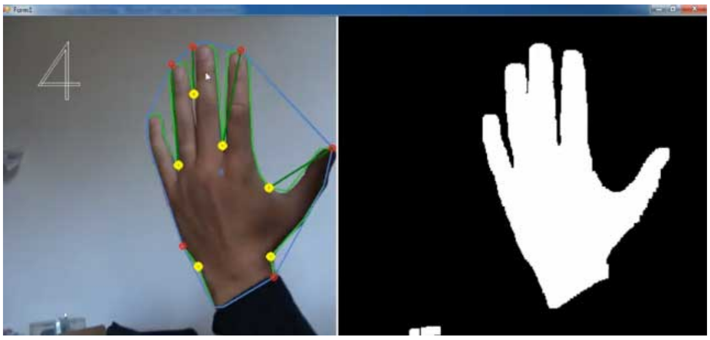


Figure 3.1 - Hand contour detection with openCV

3.1.5. Interface

Working with the limitations imposed by the technology, it wasn’t practical to develop an interface based on pointers, like we would if developing an interface for a mouse based interaction. Given that the hands of the user occupied most of the capture frame of the camera, it didn’t leave much room for that kind of interaction. Also the hand-eye coordination demand would be too high for a PAD interface setup, were the interaction time of each user is normally very small, and a frustrating user interface would reduce that span even further.

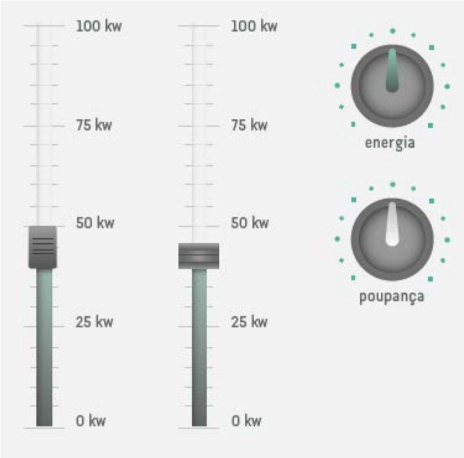


Figure 3.3 – Interface Knobs

The interaction was developed by reproducing fine hand gestures that involve mainly finger and wrist movement, in order to adjust a set of control knobs of different shapes and purposes (Figure 3.3). The user would be presented with a scenario that would be fine-tuned using a set of knobs and switches. A possible scenario would be a house where the objective could be to optimize the power consumption. To do that, the user is presented with a set of different knobs, for example: one to control the number of light bulbs that are on; another to control the level of a thermostat and another controlling the number of electrical appliances connected to the power grid. In this scenario, the user would have to adjust the knobs to save as much energy as possible, but also maintaining the house comfortable and functional: the number of light bulbs should be reduced but not to the point where the house is too dark; the thermostat should be reduced, but maintaining the house with an acceptable and comfortable temperature. Each knob would have a different gesture associated to it, like a vertical or horizontal sweeping motion for a sliding knob (left on Figure 3.3), or a rotating motion (right on Figure 3.3). The motions are distinct so that the user does not have to select the knob that he wants to use, in a pointer-like action, he would only have to perform the gesture associated with a particular knob.

Figure 3.2 - Gages

The information would be presented to the user through a set of gages (Figure 3.2) that would translate the action taken by the user into measures that he could relate to, like the amount of money that he would save in a year by applying the changes to the scenario, the amount of greenhouse gases that would not be emitted or the amount of energy being used in the scenario. The application would take the role of medium, leading the user to test different variations of the presented scenario, and possibly persuading the user to implement the same behaviour in his daily life, contemplating the knowledge acquired from the interaction (Fogg B. J., 2002).

Further work on the interface concept was postponed until the early testing of the technology was done.

3.1.6. Early testing

Small scale, informal tests were done to assert the technologies features and limitations. The work started with a simple application that was based on OpenCV and written in C++ with the purpose of testing the effectiveness of a basic hand contour recognition algorithm. The user would present a hand in front of the camera, and the program would draw the contour of the hand and show complementary information, like count who many finger were up, draw circles around each hand and identify hand joint points (similar to *Figure 3.1*). The setup was composed of a white background in a well light room, to assure the best working condition for the detection process. The camera used was a webcam, with a maximum resolution of 640x480 pixels.

The results were mixed, with the application working very well for some users and not so well for others. The recognition process had an irregular performance, with many constraints involved in the detection process. Simple details like glimmering finger nail polish, would cause a reflection that would be picked up by the camera, giving the application a hard time identifying the contour of the hand. Same result for users with hand accessories, like rings or wrist watches. Other phenomenon's like loss of finger tracking due to shadows produced by jacket sleeves, or even by the hand itself, also produced unacceptable accuracy. Even after tweaking the prototype's parameters, the results were still unsatisfactory. In a test scenario, outside of a laboratory environment, where the background for the camera shots was not constant, or the lighting wasn't perfect, the results could potentially get worse. The testing and development based on this technology was seized, deeming it inadequate for the implementation of the previously described prototype concept.

3.1.7. Conclusions

The technologies applied for this test stage had their limitations, but if the conditions for the prototype testing were controlled, they could be a applicable for the final PAD prototype. But given the concept defined earlier, this technology would not offer a viable solution for the development of the prototype.

Nevertheless this stage of development allowed us to experiment with gesture based interface concepts and draw a base line for future work, resulting in the first glimpse into what the final prototype could be.

3.2. Early 3D sensor based approach

3.2.1. Bridging previous concepts

In this stage the application concept presented in the previous chapter is maintained, but a new technological solution for the development is proposed. For this stage the objective was to select a new method for gesture recognition that did not have the constraints that the webcam based approach had, regarding reliability, precision and setup.

3.2.2. Technologies considered

By the time we were finishing the development of the first stage, a new device was conveniently put on the market, the Microsoft Kinect, a "controller-free gaming and entertainment experience" by Microsoft for the Xbox 360 video game platform" (Kinect, 2011) (Figure 3.4). One of the first widely available, PC enabled devices equipped with a depth sensor, RGB camera and multi-array microphone "which provide full-body 3D motion capture, facial recognition and voice recognition capabilities". The Kinect has very few limitations for the depth sensor, a distance of about 1,8 meters from the sensor to the users is needed, as well as the area around the users' needs to be as free has possible of objects like furniture for best results. There are also no lighting constraints, it can even be used in a pitch black environment, a frequent problem encountered on other vision based technologies.

Figure 3.4 - Microsoft Kinect for Xbox 360

This sensor was also chosen for this project, because of the open source support offered by the manufacturer of the 3D sensor technology, an Israeli company called PrimeSense that supplied a driver for this equipment. They also released a middleware, called NITE, for the driver that handles basic presence detection/position, supporting full body skeleton tracking (Figure 3.5), granting liberty for HCI designers to explore means of interaction. This software is based on the OpenNI framework specification, "which provides

Figure 3.5 - NITE Skeleton Tracking

an application programming interface (API) for writing applications utilizing natural interaction" .This framework is based on the necessity to "certify and promote the compatibility and interoperability of Natural Interaction (NI) devices, applications and middleware".

NITE does not fully implement the OpenNI specification, since it does not support the tracking of all the main body joints (Figure 3.6), like the finger joints of the hands. The main reason for this is the low resolution of the device, the 3D camera has a resolution of only 640 by 480 pixels. Given that the user has to stand 2 meters from the sensor, the stream quality is not enough to reliably support a great number of joint tracking.

In order to bridge the previous prototype concept, a mean of extracting the missing joint points of the hands was needed. To fill this need, a new technology was tested, a Kinect plugin was being developed to integrate MIT's Robot Operative System(ROS), that allowed us to process the extracted Kinect point cloud, together with the NITE joint information, to obtain the needed information. The ROS supplies a virtual environment that allows us to reproduce, even graphically, the captured live action of the Kinect (Figure 3.6). The main constraint identified during testing was the short distance to sensor that the hands needed to be placed, in order to extract enough information to be able to consistently calculate and track the position of all the fingers. The user needs to produce the gestures at a distance of 0.6 to 1 meter from the device, and the hands would occupy most of the captured frame. Since this particular constraint would fit the needs for the presented concept, the development proceeded.

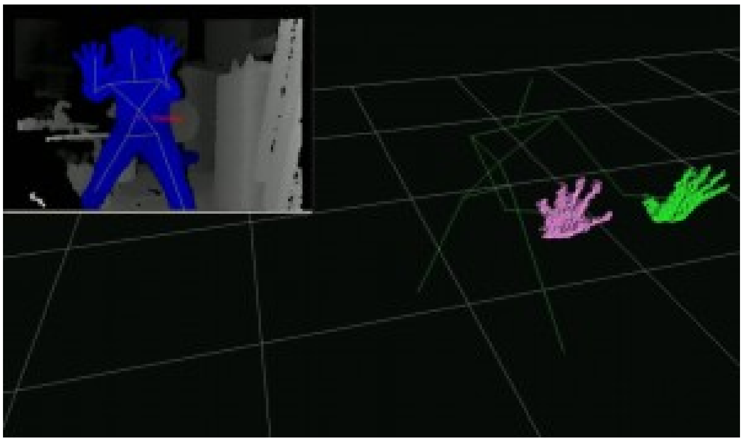


Figure 3.6 - ROS hand detection

3.2.3. Proposed Technological Solution

The main element of the architecture of this prototype is the Kinect. The feed of the camera is processed by the NITE middleware, which extracts the main skeleton joints and the full 3D point cloud of the feed. This information is processed through the ROS framework, via the Kinect plugin, used to extract precise hand positioning. The hand tracking results are used to identify hand gestures that are applied to the PAD interface.

3.2.4. Interface

The interface concept remains the same as the previous prototype concept, with an added element: a calibration phase. This technology needs the user to perform a calibration gesture, so that the NITE middleware can identify the user, and calibrate the dimensions and proportions of the user's body. To calibrate the full body joint detection, the user needs to stand in front of the sensor with his hands in the air, flexing the elbows in a 90º angle and with his legs apart (Figure 3.7). The calibration phase precedes the main interface phase, and needs to be repeated each time the user is lost and tracking stops, a limitation that further interface designs have to take in consideration.

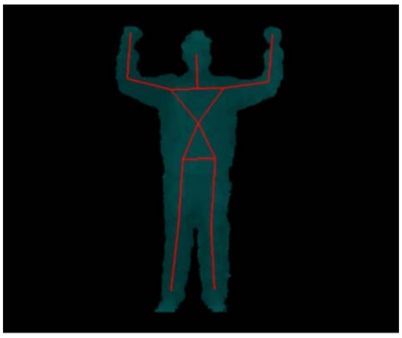


Figure 3.7 - NITE Skeleton calibration position

3.2.5. Early testing: KPong

The first test phase was dedicated to evaluate the reliability of the basic Kinect technology and the NITE middleware, avoiding in this early phase the complexity layer that the ROS framework would add. The main points that were addressed were: How many users could the system support? How much effort was needed to calibrate and maintain the joint tracking system? And how accurate and reliable was the detection?

To answer these questions a test application was created based on the hand detection demo that was supplied by the NITE setup. Inspired by the use of simple computer games to test new HCI platforms (Eisenhauer, Lorenz, & Zimmermann, 2008), we reinvented a classic video game called PONG, but instead of the original analogue controllers, we used hand tracking for controlling the paddles instead of the classic handheld input device. Pong is a crude ping-pong simulator created in the early 1970's for the arcade market. It consisted of a game environment that resembled a ping-pong table viewed from the top, with a paddle on each side, represented by two straight vertical lines (Figure 3.8). The gameplay was limited to each player moving the paddle up or down to make contact with the ball, and depending on the point of impact with the paddle, the ball would invert its trajectory with a specific angle and it would accelerate with each hit. The player would win a point for each time the ball passed by the paddle line of the opponent. Our implementation,

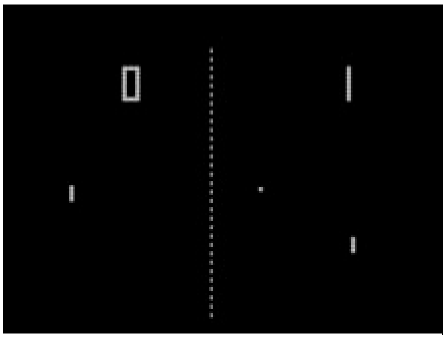


Figure 3.8 - The Original Pong

that we called KPong, resembled the original concept, with the same game environment and gameplay. The KPong would detect if a potential user was in the field of vision and ask him to perform a calibration gesture. When that was completed it would then ask a second player to perform the calibration and then the game could start. If a player leaves the camera field of vision, or the hand tracking was lost, the game would pause and wait for recalibration to be complete. For this test only hand detection was used, involving less effort to calibrate (only a small hand waving gesture). An extra game play mode was also implemented, where the user could freely control the paddle in 2D along his side of the game field. This mode was added for further precision testing. Another issue that was tested was the need for a mirror-like representation of the camera feed, so the user can centre himself in the camera field of vision. KPong has a small frame, where the mirror camera feed was projected, on the top-centre area of the game environment, and could be toggled via a keyboard key.

The application was tested on an open-day event, aimed at a young crowed of school students, called ExpoFCT. In this event the Faculdade de Ciências e Tecnologias receives students from all over Portugal to show what the faculty has to offer in terms of education, allowing future college students to get a brief glimpse of each course. The setup was done in a classroom with a cleared area and an overhead mounted projector. The students arrived in groups, and were paired to play the game, so there was always an audience in the field of vision of the device. The event supplied a steady stream of users to interact with this installation, with ages ranging from 4 to 17, playing the game in one-on-one matches, with some users interacting with more ease then others, but the reactions were always positive. The calibration of multiple hands was one of the main problems. The first hand calibrated easily, while the second sometimes took a little more effort, resulting in a frustrated user franticly waving his hand in front of the sensor. But once the hands were calibrated, as long as the players kept their hands in range of the camera, there were no problems. Normally most of the situations where the hand tracking was lost occurred when players rested their hands by lowering them and keeping them against their bodies for a couple of seconds. The users normally had a good sense of

Figure 3.9 - KPong at ExpoFCT in 2D mode

the camera's field of vision, so the mirror frame was only toggled-on in the calibration stage, for the users to grasp the concept of the interface more easily. When the game started it was normally switched-off. Since the interaction time was limited to five minutes, we had many eager players waiting in line, including even the teachers that accompanied the students.

The 2D paddle controlling mode (Figure 3.10) posed a different challenge to the players, it required more accuracy, but surprisingly that factor ended up not being the most relevant, compared with the challenge of maintaining the players hand in reach of the device.

Figure 3.10 - KPong in 2D paddle movement mode

With the new challenge presented, the players hurrying to race the paddle across the screen to reach the ball did not pose any difficulty to the tracking system, unless the hands were out of sight of the Kinect. With the pressure of the game, this was frequent, but after a couple of times having to recalibrate the hand tracking, the players would learn to keep their hands within the game area.

The hand tracking revealed to be accurate up to 6 meters from the capture device, so the reliability and accuracy was enough to suit the needs of the project. The calibration of the first players' hand was easy for this case, but the second player revealed to be more complicated and time consuming, requiring assistance from the supervisors most of the times. It was also difficult in more challenging gameplay settings to keep the users from reaching beyond the field of vision of the camera, thus losing hand tracking. During this day of testing, body types and sizes, with different types of clothing seemed not to take relevant effect on the recognition accuracy

3.2.6. Early testing: ROS

After the success of KPong, it was time to further test the target architecture, which involved testing the ROS framework and the Kinect plugin for it. The test started using the demos that the setup supplied, that allowed full hand detection (Figure 3.6). This application used the skeleton detection of the NITE middleware (Figure 3.7). This revealed to be a major constraint, since the calibration was not as easy to explain as the hand tracking calibration. It had a low success rate and it consumed a lot of space to reproduce. In order to use the ROS based hand detection, after the user calibrated the skeleton tracking, he then had to present his hand fairly close to the sensor for it to detect the hands.

After some further testing, the combination of the ROS framework and the Kinect plugin was deemed very unstable, with frequent crashes. Despite these flaws, the hand detection was very accurate, leaving hope for a future version of the software that addressed the mentioned issues. No further testing was done with this technology.

3.2.7. Conclusions

The stability issues of the ROS framework, combined with the unpractical and frustrating calibration procedure of the full skeleton tracking, lead to the conclusion that this architecture was not the ideal for reaching the goal of a flexible setup that could be used in multiple conditions, like crowded spaces. But the basic hand tracking of the NITE framework worked quite well, and would be further used in the development of the next prototype due to its easy calibration, good accuracy and reliability, proven in these early tests. But the use of this technology involved redesigning the interface concept for the prototype, from hand/finger gesture to a pointer based approach, where the position of the user’s hand is translated in to a single point position in the interface.

Given the presented findings, the system was rethought, taking in consideration the defined goals and requirements, as well as the tested technologies. This process resulted in the implementation of the prototype presented in the next chapter.

4. The Maid

Many environmental problems occur due to our erroneous and exaggerated use of Earth resources, in our endless thirst for energy. The most common energy sources are based on the burning fossil fuel, which lead to the emission of green-house gases. The search for new cleaner energy sources isn't the only battle to be fought, each one of us should also strive to reach a more efficient use of our energy resources, and that battle begins right in our homes. Everyone can contribute to reduce the domestic energy consumptions and consequently greenhouse gas emissions. Small individual actions may seem insignificant, but together everyone's efforts can have a great impact on the environment.

Figure 4.1 - MAID logo

The prototype developed in the scope of this thesis uses persuasive technologies can be used to make people aware of the consequences of their actions and teach them how to proceed, helping them change their attitudes and behaviours. Motion-based Ambient Interactive Display (MAID) (Figure 4.1) is an application deployed on an interactive public ambient display (PAD) system, which explores the use of natural hand gestures as a mean of creating a seamless human-computer interface, delivering a persuasive interface. The main objective of MAID is to instruct (or remind) users about simple procedures to save energy, showing them how easy it is to have a huge impact on the environment by taking simple actions in their daily life. MAID provides users with an interactive scenario, which can be manipulated simply by hand gestures. This scenario is a simulation of a domestic setting, a house composed by a set of rooms that the user can explore, trying to find out what could be done to improve the overall power consumption and implement the corresponding changes, therefore acting as medium so that the user can practice a target behaviour, leading to a behaviour change (Fogg B. J., 2002). The system can also take the role of a social actor, since the MAID user could have an audience, and the actions on the interface can act as social pressure, that can also influence the audiences' behaviour. This implementation is done by directly interacting with the virtual object by performing a specific gesture. The impact of the implemented actions in the game is explained by audio feedback and can also be visualized by the amount of money the user will save and the reduction in CO2 emissions, via the heads-up display present in the interface.

The gestures used to manipulate and improve the power consumption are applied in a way that resembles the corresponding real world action. For example, to pull a laptop's power adapter from the cord, a pulling gesture will be used. This not only contributes to create an engaging experience, but also establishes the connection between a simple gesture and the impact that it can have when applied to the real world.

In this chapter we will explore the MAIDs architecture, describing the interface and technologies applied and finally showing the results obtained in our user tests.

4.1.Architecture

In most common implementations of a PAD, in order to add interactive components to the interface, the interaction method used is integrated in the display itself (touch-screen), or attached to it, in the form of button panels or hand-held controllers. For the MAID, the centre piece of the architecture is a motion-detection 3D sensor device - “Kinect for Xbox 360” by Microsoft – (see 3.2.2) that is used to obtain the ambient contextual data. This data is extracted from the device using the NITE middleware that pre-processes the 3D point cloud gathered from the camera feed and supplies a basic abstraction layer for tracking of the users’ position and tracking of their hands. The user detection, since it does not require user calibration, is used to detect the presence of an audience, information used to dynamically change the interface content. The hand tracking is used for direct user interaction. Since it requires user calibration, this data is used to emulate a pointer device based on the user’s hand position and movement

The MAID development was done on Linux based operative systems, since early on most of the software development involving the Kinect was only available on that platform, so most of the technology options for the development process were based on what the Linux platform had to offer.

The prototype is built over the NITE middleware, using the FreeGlut implementation of the OpenGL framework for graphics processing, being one of the most widely supported 3D graphic platform available for Linux. The textures applied in the application are loaded from png files, interpreted by libpng, a Linux png processing library. The png file format was used due to its compact format and transparency support. For configuration file management the development was supported by the libconfig Linux library, following a highly modular and configurable development strategy in order to allow gameplay scripting. Most of the components of the interface are configurable via a configuration file, like all the textures used for the graphic rendering, the size and type of the fonts applied, sound tracks, interaction parameters, textual content, among many other options. The gameplay itself can be modified, by adding or removing rooms to the scenario and adding or removing interactive elements from each room.

The selected programing language for the implementation of the prototype was C++, mainly because of the NITE middleware, for there were no wrappers for other programming languages. Since C++ is the common thread of all the selected technologies, that did not pose any difficulties. It is also one of the most used programing languages.

Another component of the MAID is a sound engine that serves as a complementary mean of communication. Since a public setting is full of potential distractions, a PAD should resort to more than one of the user's senses. This module reads out loud most of the textual information presented in the interface, via Festival text-to-speech engine, a popular framework available for Linux. It also manages the playback of the application soundtracks using the FMod framework, a cross-platform system that supports multichannel direct playback of multiple audio formats, like wav and mp3. The implemented MAID architecture is presented on Figure 4.2. The main components of the architecture are the following:

- **Control Centre:** This is the centre piece of the architecture. It instantiates the main modules of the structure, based on the configuration data. When this module is instantiated, it starts by initializing the data acquisition modules: The configuration module and the NITE Data Acquisition module. It instantiates the Audio Manager, Interface Manager and initializes the OpenGL Frame Drawing loop and update loop. It also starts monitoring the keyboard key presses, for basic commands, like toggling full screen mode, pause the interface, toggle mute mode, reset the game and toggle Debug information displaying. This module holds basic control over the application by receiving the processed NITE information and redirects it to the interaction management module, together with the request to draw the interface frames.
- **NITE Data Acquisition:** This module is part of the Data acquisition layer. It establishes a connection to the NITE middleware, which processes the point cloud collected from the Kinect. When it is initialized, it opens two NITE tracking clients, one to monitor actor presence and position and another to calibrate and track the hand of a user. The hand tracking client information is only processed when there are actors present. The hand and actor information is updated with each rendered frame, matching the applications frame rate. The NITE sends the information in the form of a 3D coordinate array, in each frame, the actor tracking data is updated and if there are actors present, the hand tracking data is updated.

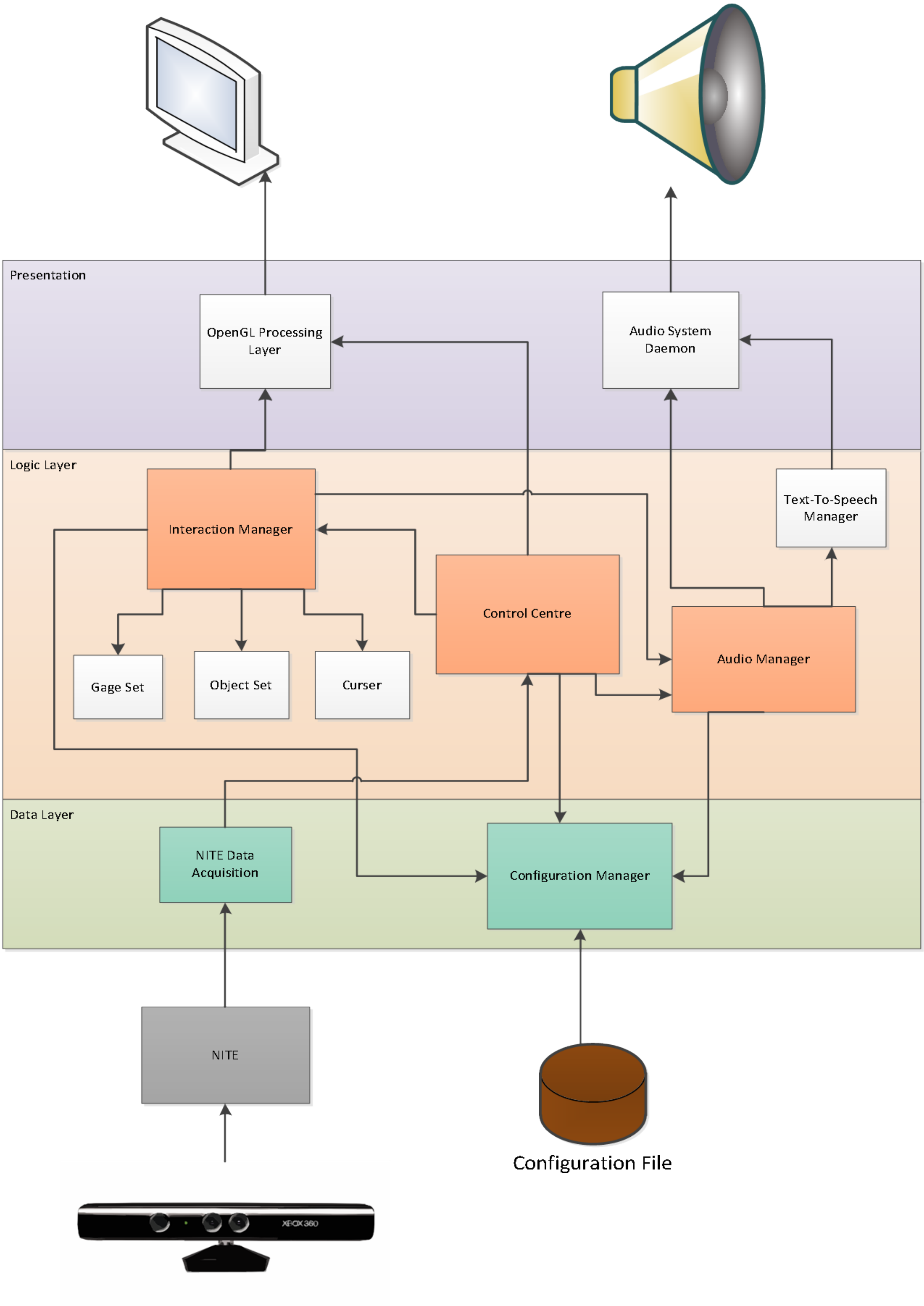


Figure 4.2 - MAID Architecture

- **Configuration Manager:** This module has the purpose of interpreting the configuration file by using libconfig, a configuration file parser available for Linux based operative systems. When the module is initialized, it uses the library to parse the configuration file, and creates the configuration option graph structure to the memory. The file must be located next to the application executable, and be named "maid.conf". The other connecting modules can then access the information, querying for the option name, or option graph sequence. If the configuration file is modified, the application has to be restarted to apply the changes. The configuration file format is specific to libconfig, as it can be verified in this example:

```
version = "1.0";

application:
{
  version = "1.0";

  window:
  {
    title = "My Application";
    size = { w = 640; h = 480; };
    pos = { x = 350; y = 250; };
  };
}
```

  In this example, the version of the configuration file is "1.0", it describes an application that has a version and a window. The window has a title, a size and a position. To retrieve, for example, the x coordinate of the position of the window, the configuration module has to be queried for the graph sequence "application.window.pos.x".

  This configuration method is used to describe most of the interface properties, for example we could configure the background textures to apply to each stage of the interaction, the sound effects and background music that will be used or even timing details associated with each action and transition.

  The complete scripting of the gameplay is done though the configuration file. Allowing the administrator to configure what objects will be available and all their properties, the different scenarios that will be available and objects present in each one, the gages that will supply the necessary feedback to the user, even the aspect of the cursor and its properties.

- **Interaction Manager:** This module manages the interface drawing, the interaction stage transition (see 0) and it also maintains the game state. When initialized, this module loads the set of interactive objects from the configuration file, including their position, the textures used for each stage of the object interaction, gesture interaction type/parameters and text information to be displayed and spoken. It also loads the set of gages presented in the game's "heads-up display" (HUD) (see 4.2.6), also loading their textures, position and parameters. It additionally loads the parameters regarding the interface cursor and stage specific elements, like backgrounds and audio soundtracks. This module manages the interaction stage drawing and transition, in a sequential way, following a set of predetermined rules to progress or regress stage. To do so, it receives, through the control centre module, the actor and hand related information from the NITE Data Manager. The transition from the first stage to the second is based on the user's presence. If a user is detected in front of the sensor, the module triggers the transitions to the next stage. The second stage transition is based on actor position, if an actor maintains its position, within a configurable range of motion, in front of the sensor for a configurable amount of time, then the next stage is triggered; if the actor leaves the sensors field of view, the interface return to the first stage. The third stage is based on actor position and the calibration of a hand. If the user maintains its position and calibrates the hand tracking, by waving is hand in front of his body, then the next stage is triggered. If the actor does not hold is position, the interface return to the second stage. The fourth stage is the final stage, it will return to the previous stage if the hand tracking is lost. At this stage the game environment is drawn, which includes a HUD and a game terrain. The game terrain can be explored by means of a cursor, controlled by the position of the tracked hand. In the terrain the user can find the interactable objects. Each stage, except for the first stage, has a soundtrack associated. So in each stage transition, this module communicates with the audio manager to change the sound track, some events can also trigger a sound effect. It also communicates with the audio manager to push texts to be spoken by the Text-To-Speech Manager.

- **Audio Manager:** This module manages all sound events of the application, from sound track and effects playback, to text-to-speech processes. On initialization, this module loads basic parameters from the configuration file, part of them are to configure the FMod engine and the other to configure the festival system. The FMod engine is a flexible, multi-channel sound engine that supplies basic abstraction to play sounds, allowing the application to manage playing mp3 files simultaneously, including sound level manipulation. This module accepts instructions to play audio files, in mp3 format, based on their path, but it does not use caching for the audio files played. File caching is implemented for sound effects by the FMod framework, so no further optimizations are deemed necessary. This module also supports text-to-speech (TTS) functionalities, using the Festival system. When a TTS instruction is accepted, the audio manager creates a new thread, if no other TTS job is running, to launch the Festival TTS operation that outputs the resulting speech directly to the Audio System Daemon. If a TTS job is already running, it cancels this previous job and reuses the available thread to receive the new job. The thread management and communication process is done using the boost library, and it is needed mainly because the Festival TTS process will block the MAID execution while it is being used if it is not running on a different thread. Festival is also very CPU demanding, so by using a threading technique, the module can take advantage of the multi-threading and/or multi-core capabilities of the CPU. The TTS process also has disadvantage of API not having a speech cancelation instruction, so once the speech is started, so the text has to be received by the TTS thread in an array of segments. The text segmentation is defined by the administrator in the configuration file, so that if the application has to cancel the speech, it doesn't cut off immediately in the middle of a sentence, the administrator can divide the speech into short sentences in order to keep the gameplay smooth. These segments are pooled and, if needed, the speech can be cancelled by empting the pool, and the thread can be reused by pooling a new speech.

The implemented architecture is a configurable and modular framework that enables MAID installations to adapt to each setup context. The established framework can also be repurposed by creating different gameplay scripts that could be targeted to different interaction studies, with themes and goals beyond the scope of this thesis.

4.2. Interface

The MAID has a four stage interface (Vogel & Balakrishnan, 2004), in order to supply a public ambient interface that persuades potential users to engage in a direct interaction setting. The interface is intended to blend in with its surroundings, changing its content based on the audience and its behaviour:

4.2.1. Stage 1 - Idle

Figure 4.3 - MAID Stage 1 composed with a webcam output pointed at the user

In this first stage, the MAID system is in an idle state because there are no users detected by the sensor in the surrounding area. The interface presents minimal content, like a simple logo (Figure 4.3), with no audio playback, blending in with the surroundings and not trying to attract attention, since it does not detect an audience to target. The next stage is activated when an actor is detected in the sensors field of vision and stage 1 phases out.

4.2.2. Stage 2 - Greeting

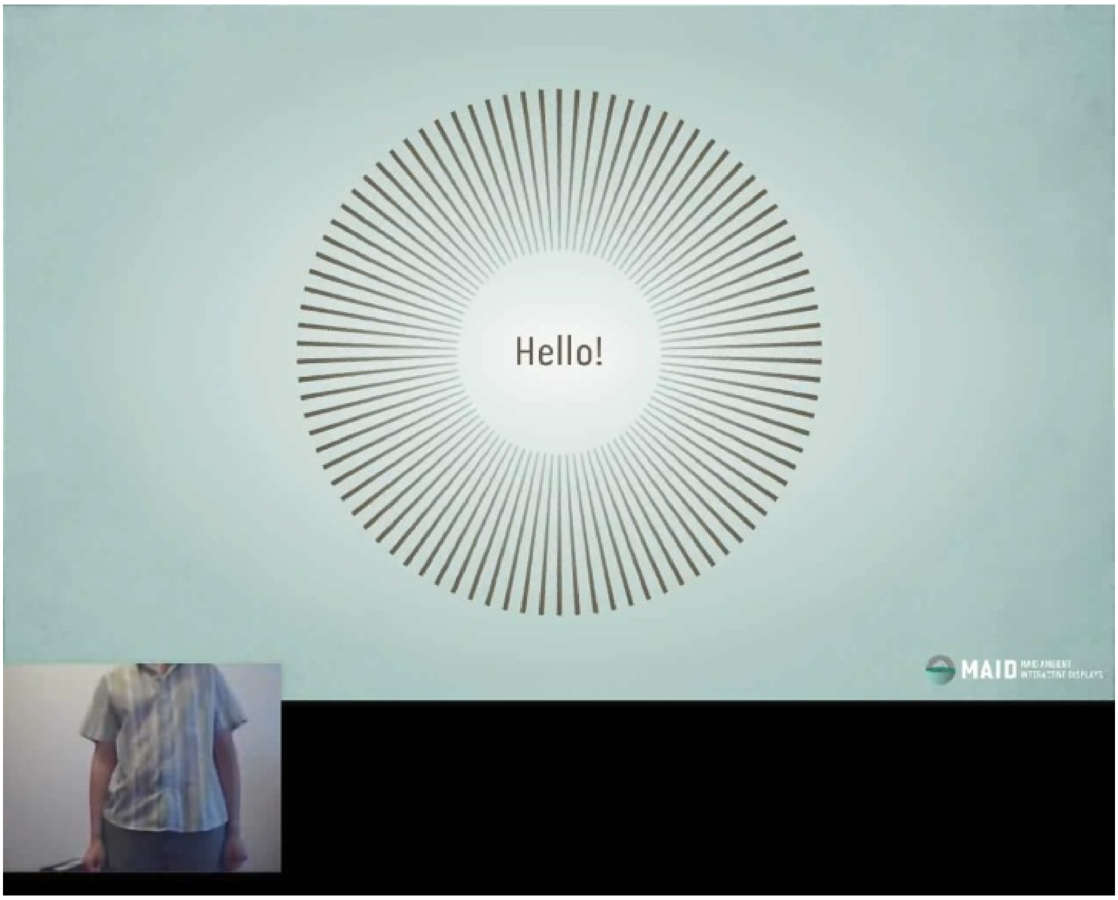


Figure 4.4 - MAID Stage 2 composed with a webcam output pointed at the user

The stage 2 interface phase in when a user is detected, so the interface tries to captivate his attention through a greeting message inside an animated circle (Figure 4.4) and can also maintain the audiences' attention by displaying persuasive content in the background. In this example (Figure 4.4) no special persuasive content is configured, so the default background is shown. The content to be displayed in background is intended to be generic, regarding the applied theme. The MAID tries to capture the attention of the user through audio, by playing a soundtrack, pre-configured for this stage, and reading a welcome message. The background content can also be specified in the setup. If the MAID maintains an audience for a configurable amount of time, it moves on to the next stage, concluding that the users are interested in the content that will be displayed.

4.2.3. Stage 3 - Calibration

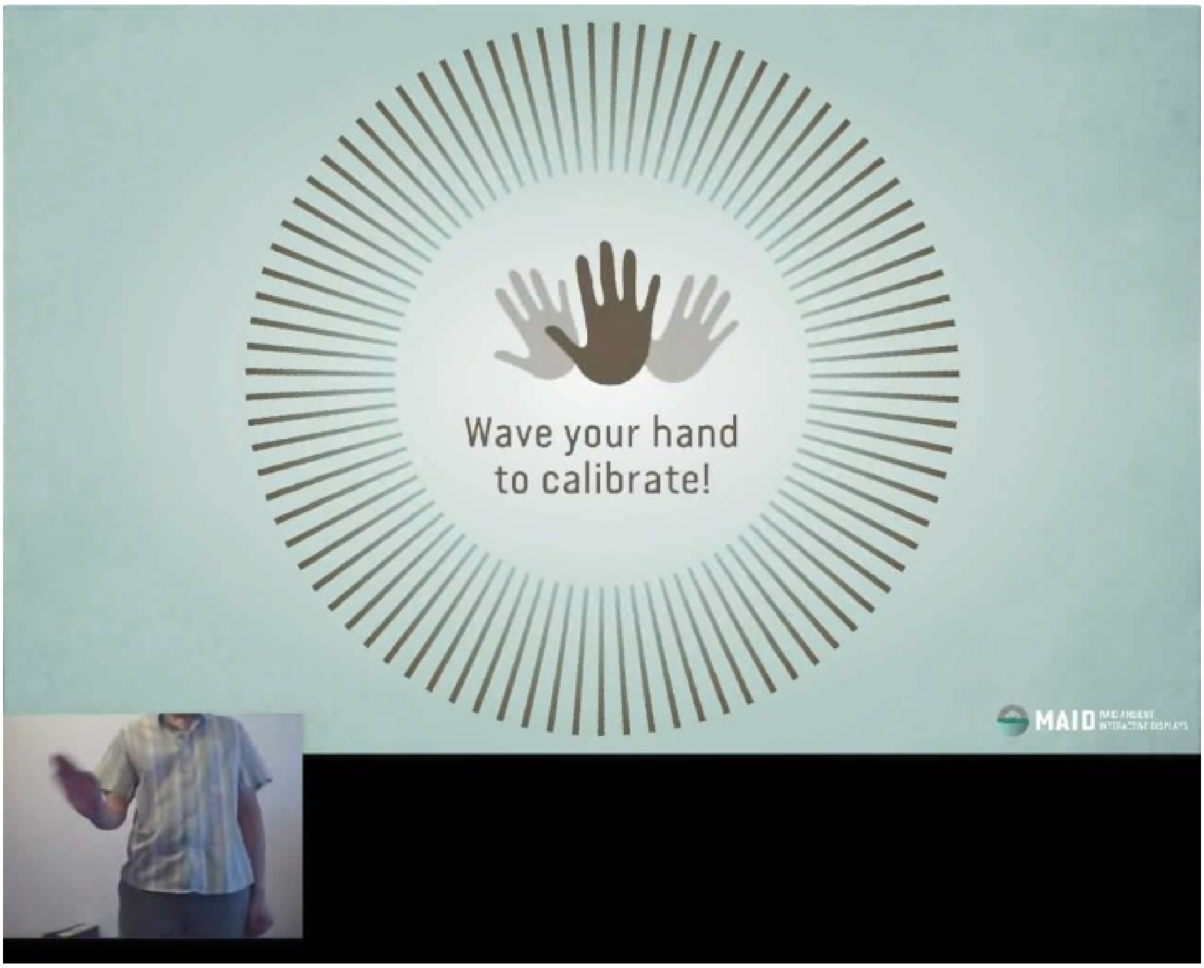


Figure 4.5 - MAID Stage 3 composed with a webcam output pointed at the user

An actor has been identified as a captive audience and is invited to further interact with the system. The actor is asked to produce a wave-like activation gesture with his hand in order to calibrate the gesture recognition system. This stage is used to identify if there is a user among the audience who wishes to proceed interacting with the system. The background information is intended to introduce the user to the stage 4 interface, persuading him to calibrate the hand tracking mechanism. When a user of the audience calibrates his hand, the next stage is triggered.

4.2.4. Stage 4 - MAID Home Screen

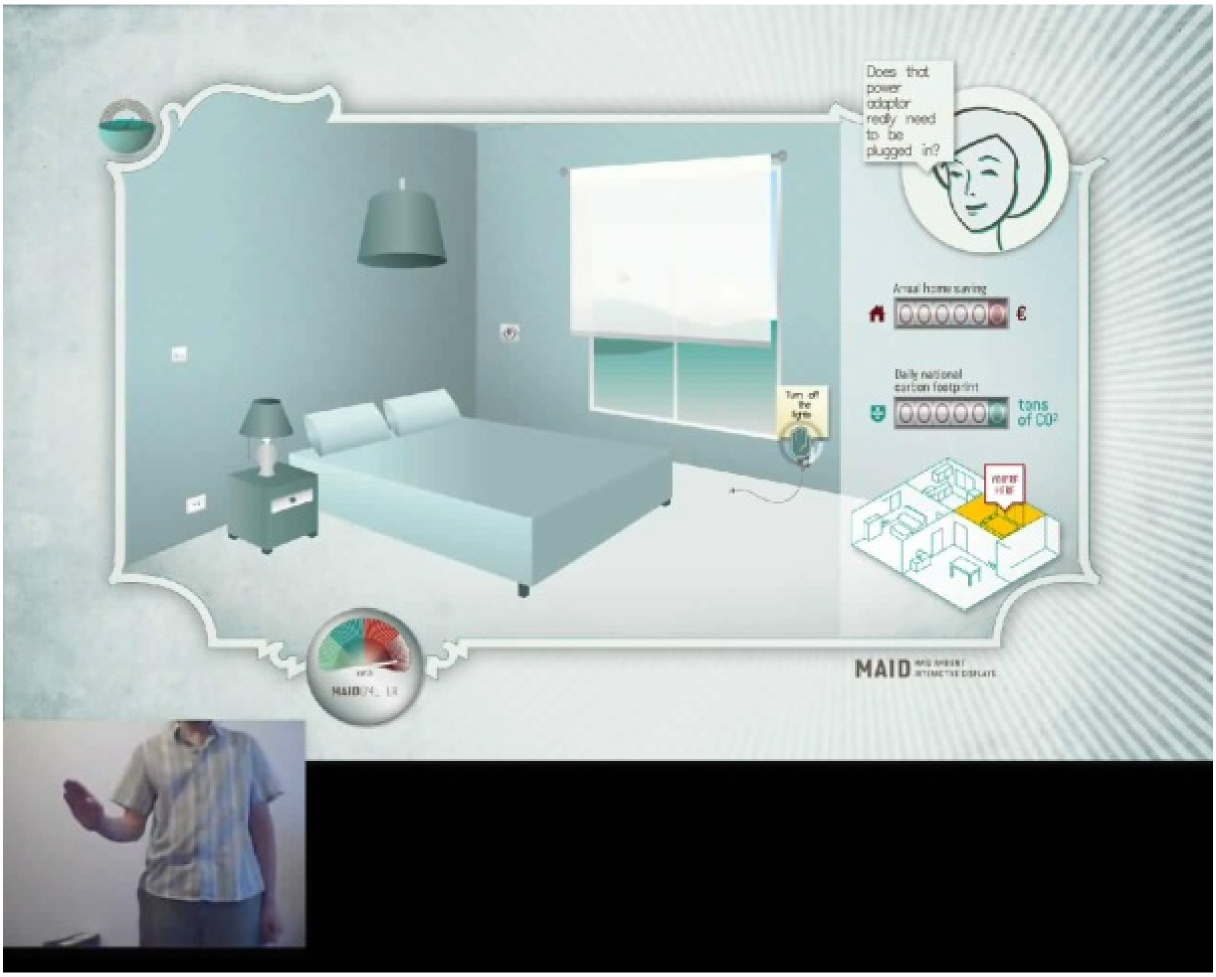


Figure 4.6 - MAID Stage 4 composed with a webcam output pointed at the user

In this stage the user is interacting with the main content of MAID. The system also supports a situation where there is an audience behind the user without negatively affecting the hand tracking process. An interactive scenario is presented to the user. He can now explore it through hand gesture recognition.

The first three stages are articulated in order to build momentum and anticipate the final stage. At this point, a typical video game interface format is presented, composed of a head-up display (HUD), a game terrain and a pointer (Figure 4.6). This interface stage also features a direct interaction mode, which allows users to perform actions over selected objects of the virtual environment, like changing a light bulb on a lamp, by using hand gestures. If the hand tracking is lost, the interface returns to the previous stage, to recalibrate. To exit this stage of interaction, the user just has to walk away from the PAD. As soon as the MAID identifies that the user is no longer there, it goes back to stage 1. All the game information related to the previous user is resected.

4.2.5. Game Terrain

The game terrain is a 2D representation of a scenario. For this demonstration the terrain represents a common household, with a bedroom and a kitchen for the user to explore (Figure 4.6). The user will start in a room of the house, and at any given time, he can select another room. Each room has a set of interactive areas/objects, in a gameplay model similar to what could be found in a point-and-click adventure game. These areas correspond to situations that should be corrected, representing desirable user behaviour changes. The set of interactable areas is configurable, allowing the system administrator to define new areas, based on the current predefined gesture and interaction set specified. For example, a new interactable area that represents a TV set on standby could be implemented and a pushing motion could be used to turn it off and save power. Each interactable object has a tooltip associated, describing the action that can be taken and also a helpful speech that is read by the systems' TTS module.

4.2.6. HUD

The HUD is composed by a configurable set of gages presenting the current game status in a graphical representation that the player can easily assess. The MAID keeps track of the current total consumption of the scenario. Each interactive object has a configurable value for its consumption before and after it was optimized. Based on the sum of the current consumption of all the objects, each gage does its own extrapolation of that value, to reach its goal. The set of gages is fully configurable, allowing us to define new gages of two specific types (Needle gage or split-flap display), their design, position, size, scale and even the sound effects that are to be applied. The set of gages configured for the latest demonstrations include:

- A classic needle gage showing the current consumption of the home in Kw/h (Figure 4.7). The gage scale, instead of showing numeric values, is divided in four quadrants, and the quadrant that corresponds to the current position of the needle shines brighter, using a graphical representation of the values. The needle itself is also constantly vibrating too further emphasize the importance of that area of the interface and call the users' attention. When the house consumption is lowered, the needle slowly descends, contributing to a gratifying dramatic effect. The gage quadrants implement a color code, in the far

Figure 4.7 - Needle Gage

right the tones are redder, representing a bad current score, and to the far left corner the tones are greener, representing a good score.

- A counter gage, presenting the daily CO_2 emissions that would be avoided if the actions currently taken by the users in the game were also taken in every home in the country. For our demonstration it was configured for the population of Portugal, in t/m2 (Figure 4.8). This counter gage is animated to resemble a split-flap display, so when the daily CO_2 emission calculation goes down, the number on the counter starts to decrease rapidly, slowing down as it gets closer and closer to the target value. It also has a clicking sound effect to call the users' and audiences' attention to the impact of their actions.

Figure 4.8 - Counter Gage

- A counter gage, showing the home costs saved if the actions currently taken by the users in the game were taken during the course of a year, in euros. It is animated just like the previous counter gage. These gages help to deliver the persuasive message, urging the users to realize that small individual actions can have a significant impact in the environment. This is done by presenting the user with actual values that the user can relate to and understand their impact.
- In the HUD there is also an interactive house plan that allows the user to move, from one room to another, in the house (Figure 4.9). The user can select a room by hovering the cursor over the areas of the floor plan. The number of rooms present in the game script is configurable, with each room representing a different scenario, with its own set of interactable objects. The total number of rooms and objects that are configured help establish the maximum time span of the user interaction.

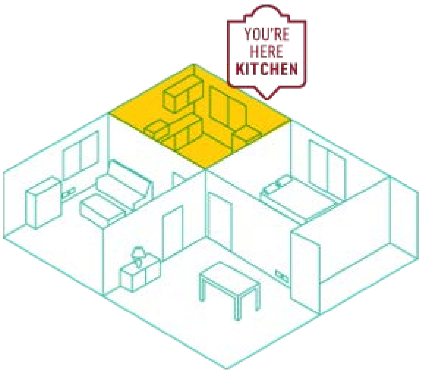


Figure 4.9 - House floor plan

- The final component is the MAID mascot that delivers information to the user when the context demands so, in the form of a pop-up speech balloon that is also read through a text-to-speech system. The mascot gives instructions that help the users to complete tasks and hints leading the users to accomplish the objectives of the game. The messages that the mascot shows in text are also spoken by the TTS system, giving the impression that MAID's speaking voice comes from the mascot.

4.2.7. Pointer

The pointer is directly controlled by the hand of the user during the exploration of the 2D scenarios. The users' hand coordinate, detected by the sensor, are translated to the 2D game terrain, so that if the user stands to far from the sensor he will be forced stretch is arms to reach the far edges of the interface with the pointer. This way the user is persuaded to take a position closer to the sensor, in order to make smaller hand gestures, therefore needing less space to use the interface and also being in a better position to read the displayed contents. In each scenario there is a set of hoverable areas that correspond to objects that the user can interact with. When the pointer hovers over one of these areas, a circular progress bar pops up, representing the hovering elapsed time. While hovering, a tooltip pops up with information related to the actions that could be taken to optimize that particular situation. The MAID mascot, in the HUD, also delivers a speech regarding the hovered area. If the hovering lasts for the full load time, the user then enters the direct interaction mode. This pointer hover procedure is also the same for changing rooms in the huds' floor plan.

4.2.8. Direct Interaction Mode

Figure 4.10 - Maid direct interaction mode

When entering this mode, the interface changes: the HUD disappears, and the centre stage is cleared to house an enlarged version of the chosen object (Figure 4.10). Here the user is asked to perform a specific hand gesture, which triggers a reaction on the object.

For example, the user could enter the direct interaction mode by selecting a Thermostat in the scenario, he would then be asked to perform a circular gesture to turn off the thermostat (Figure 4.10). The correct gesture is illustrated by the ever present MAID mascot and it is intended to match the gesture that should be done to perform the corresponding action in the real world, she also describes the gesture through a small speech. Upon the completion of the challenge, a congratulatory message is displayed and spoken, along with complementary information, like fun facts, related to the action performed. Finally the interface returns to its previous state, allowing the user to explore another interactive area/object. Once the user entered this direct interaction, he had to complete the specified task in order to return to the game terrain to select another object. If the object in focus was the last object that the user had available to optimize, then the game would present a final congratulation screen that showed the results of the completed tasks.

4.2.9. Facebook Interaction

At any given time, while exploring the game terrain, the user can publish a picture with the current game status in the Facebook social network, sharing them with the world. The interface that guides the user through this process is based on the direct interaction mode, through a hoverable Facebook icon integrated in the interface (Figure 4.11), outside the frame that contains the game terrain. The icon and its placing are configurable. After hovering over the icon the interface moves to a mode similar to the direct interaction mode, where a mirror image of the Kinect RGB camera feed is presented, with a circular progress bar painted over. While the circle fills in the user has time to strike a pose and get ready for the shot. As soon as the circle is fulfilled (Figure 4.12), a picture is taken and the user is informed that the picture was published in the project DEAP's Facebook wall. In order to send the information to the social network, a Facebook Application was created, and together with the association of the application with the DEAP project Facebook page, access tokens were generated to allow us to publish the pictures directly from the MAID to the DEAP Facebook page photo gallery.

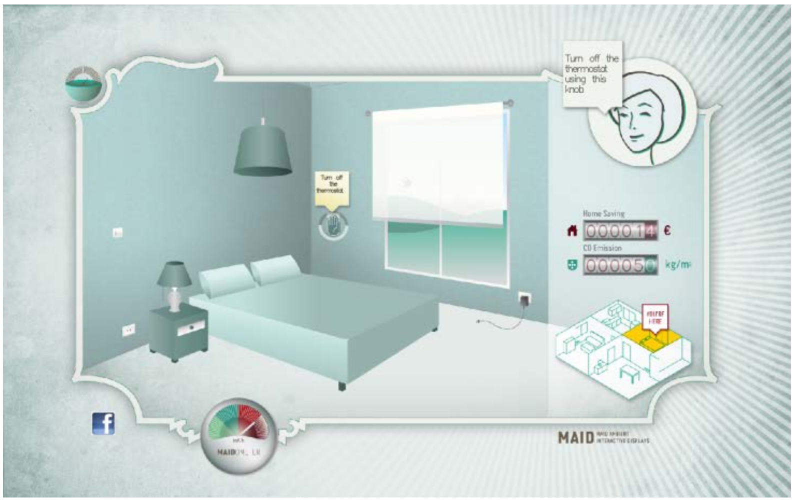


Figure 4.11 - The MAID's main interface.

This mechanism was implemented using a cURL command that sends the pictures and the corresponding descriptions to the photo gallery, using the generated access token for authentication. The picture is published in the MAID gallery of the DEAP project Facebook page, with a description that reflects the player's current progress in the game, for example (Figure 4.12):

"Action-shot taken by the MAID: At the time of this picture, the user had optimized the power consumption of the MAID home, in order to save up to 64Kw per month."

Figure 4.12 - Facebook picture capturing interface

Figure 4.13 - MAID photo gallery picture.

In this case the user had optimized the MAID home in order to save 64Kw per month. After this data is published, the result isn't presented to the user, in order to persuade him to visit the project's Facebook page (Figure 4.13), where he can "like" or even "tag" any picture, exploring the social network graph structure to spread the information and divulge the message behind the MAID.

4.2.10. Implemented gesture set for direct interaction mode

There are four distinct gesture types that were implemented in MAID and can be used to create new interactive objects. The gesture recognition is based on single hand position detection. The length of the hand movements is configurable, but generally it should not go beyond an arm's reach, so the user doesn't have to move around to produce the target gestures. This is done to reduce the complexity of the gestures, to reduce the clutter free area needed for the detection and also to increase the reliability of the hand position recognition.

The following gesture types were implemented:

- Pushing or pulling hand motion (Figure 4.14), used for objects' manipulation such as pulling a power adapter from a wall socket or tapping a light switch.

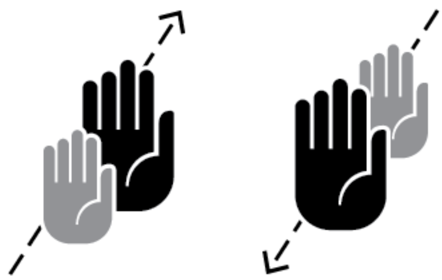


Figure 4.14 - Pushing and/or pushing motion

- Vertical sweeping hand motion (Figure 4.15), used for objects' manipulation, such as opening window shutters.

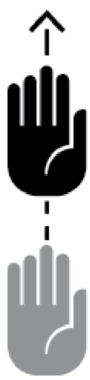


Figure 4.15 - Vertical sweeping motion

- Horizontal sweeping hand motion (Figure 4.16), used for objects' manipulation, such as opening sliding windows.

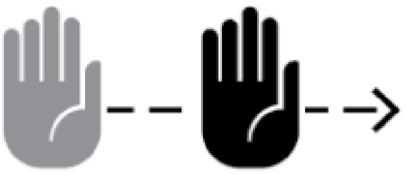


Figure 4.16 - Horizontal sweeping motion

- Circular hand motion (Figure 4.17), used for objects' manipulation, such as unscrewing a light bulb and screwing in a new one or adjusting a thermostat.

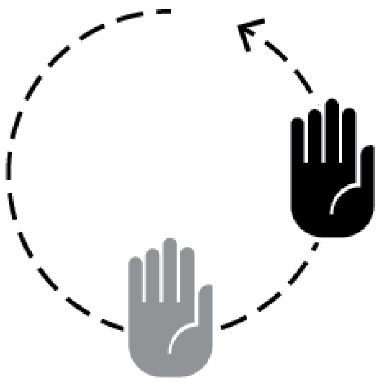


Figure 4.17 - Circular motion

The gesture commands were designed to resemble the corresponding real world actions, so they could provide a more natural interface and be easier to understand and to recall later. The MAID is also intended to work as a training system, so the users can rehearse the target behavior (Fogg B. J., 2002).

5. Evaluation

A series of user studies were conducted with the purpose of evaluating the usability of the MAID prototype, evaluate the users' reaction to the interface and there motivation to implement the actions rehearsed through the gameplay. This evaluation took in consideration factors like the gesture reproduction ease, entertainment value, the users' understanding of the action/consequence dynamic presented correlation between the proposed gestures and the action produced and dissemination by the user of the transmitted information.

For the user study preparation, MAID has three basic factors to take into account for the setup process: space, lighting and audience. Regarding space, the Kinect needs the player to be at least 1.8m away from the sensor and a 1.8m wide area around the player that is clutter free (Microsoft, 2012). Concerning the audience, the display system used has to take into account the size of the potential audience, so for a large audience we use a projector, but for a small audience we can use a large wide-screen display. Regarding lighting, regular projectors need low lighting to deliver an acceptable definition and also the Kinect sensor should not be exposed to direct sunlight.

This chapter covers the user studies done during the development of the MAID.

5.1. First User study

The first study was done in a public setting, conducted on 18 users, 22,2% male and 77,8% female, with ages ranging from 20 to 45, average age of 29 (standard deviation of 5.7). The setup was prepared in the Faculdade de Ciências e Technologia's (FCT) library. This location was chosen due to the adequate flow of potential users of its main hall, delivering a steady pace of users. Since this was the first user evaluation, it would be ideal to have a user rate that wouldn't lead to a queue of users, allowing time for eventual tweaking, tinkering and problem solving. The configuration file system of MAID was created to achieve this "on-the-spot" tweaking to adjust the content to target setup. For this evaluation study we had as much space as we needed but we had no direct control over the lighting. Since the area was exposed to natural and artificial lighting and the test should last most of the day, we opted for a display that had few lighting constraints, in this case a 32 inch LCD display, being large enough for the audience that gathered. We selected an area of the hall that did not receive direct sun light, taking in consideration the minimum the requirement of the sensor. The Kinect was mounted below the screen, and the users were positioned at a minimum distance of 2.0 meters from the sensor (Figure 5.1).

Figure 5.1 - MAID's first user study Setup

The test sessions were conducted by one researcher that, give an initial briefing and instructions to the participants and provided assistance for any problems that users might face. He also took notes regarding the user's behaviour and delivered the user assessment questionnaire.

Before starting to use the application, users were informed about the objectives of the test. After a short description of the application and an explanation of the goals to be achieved, users were encouraged to explore MAID. They could do so for as long as they liked. At the end of each test, participants were asked to answer a questionnaire to assert their reactions, gather generic personal information and collect their feedback and comments.

This version of MAID, did not yet include the Facebook interaction functionality, as it was only added and tested in later studies. The user feedback from this study would also answer the question of how open are the users to the option of publishing content on Facebook directly from MAID.

5.1.1. Questionnaire

The questionnaire captured users' personal data, experimental feedback and comments (see annex 1). It is divided in nine question groups:

1. Acquire basic personal information like the participant's age, gender,
2. Assert the degree of instruction of the user (Chart 5.1).

3. Evaluate the frequency of use of modern technologies like Computers, Smart Phones, Gaming Consoles and the Internet. Evaluating in a scale reaching from never (0) to daily (4). In this question group, the user is also asked if they ever used the MAID (Yes/No), and if they ever used the device to classify there Frequency of use from rare(1) to daily(4) .
4. Acquire general feedback regarding MAID (Chart 5.4). The users were asked to rate statements, using a five-point Likert-type scale, which ranged from strong disagreement (1) to strong agreement (5).
5. Using a five-point Likert type scale, users are asked to rate statements regarding general usability issues of MAID (Chart 5.5). It also included a sub-group of statements specifically dedicated to evaluate the appropriateness of the gestures, regarding ease of reproduction and appropriateness (Chart 5.6).
6. Assess the reception of the persuasive message by asking a set of questions, regarding power consumption and the consequences of excessive use (from Chart 5.7 to Chart 5.10).
7. Based on the Microsoft "Product Reaction Cards", users were asked to choose the words that best described their experience (Chart 5.11), while using the application from a list of 16 words, a set that included about 11 of words considered positive and 5 considered negative. The participants could choose any number of words that they deemed adequate to describe their reaction. This group refers to the reaction to users had to MAID.
8. Also Based on the Microsoft “Product Reaction Cards” (Chart 5.12), with a set 16 words, that included 9 words considered positive and 6 words considered negative. This group regards the usage of the MAID.
9. An open question, in order to gather comments and recommendations regarding future developments of additional features and to obtain a more general evaluation of the system

5.1.2. Results and Discussion

As it was previously stated, this study has a sample of 18 users with an average age of 29. The academic degrees of the users are as shown in Chart 5.1, where we can see that most of the users have a high education level.

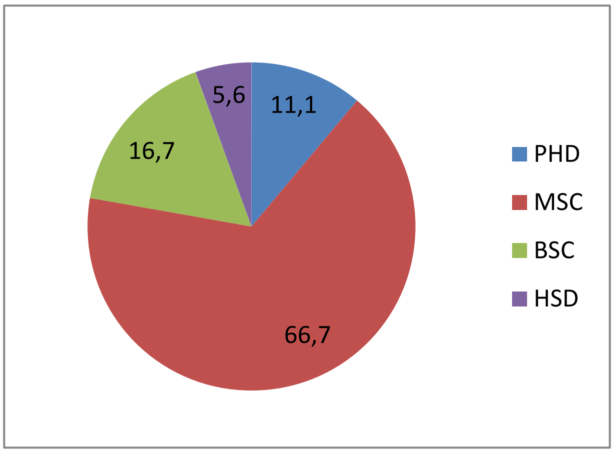


Chart 5.1 - User Academic degrees

The results of the third group of questions show that most users are no strangers to technology. This can be related with the education level of the participants. All of them use computers and internet daily and 72% of the users are daily smart phone users. Regarding game console usage, the results were less consistent as it is shown in Chart 5.2. All the participants also reported having used the Kinect sensor at least once, but only two of them are daily users (Chart 5.3).

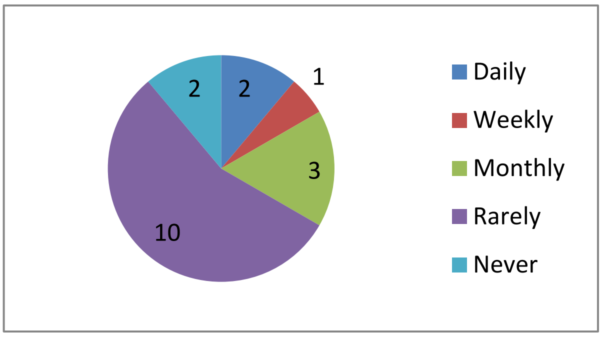


Chart 5.2 - Gaming Console usage

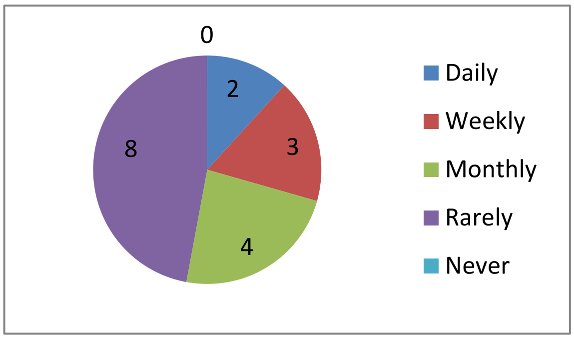


Chart 5.3 - Kinect usage

The fourth question group had consistently high results (Chart 5.4), leading to the conclusion that the users enjoyed their experience with MAID. The statement that had a marginally lower average score was “it’s easy to use MAID”, this is consistent with the remaining results of the questionnaire that will be addressed next.

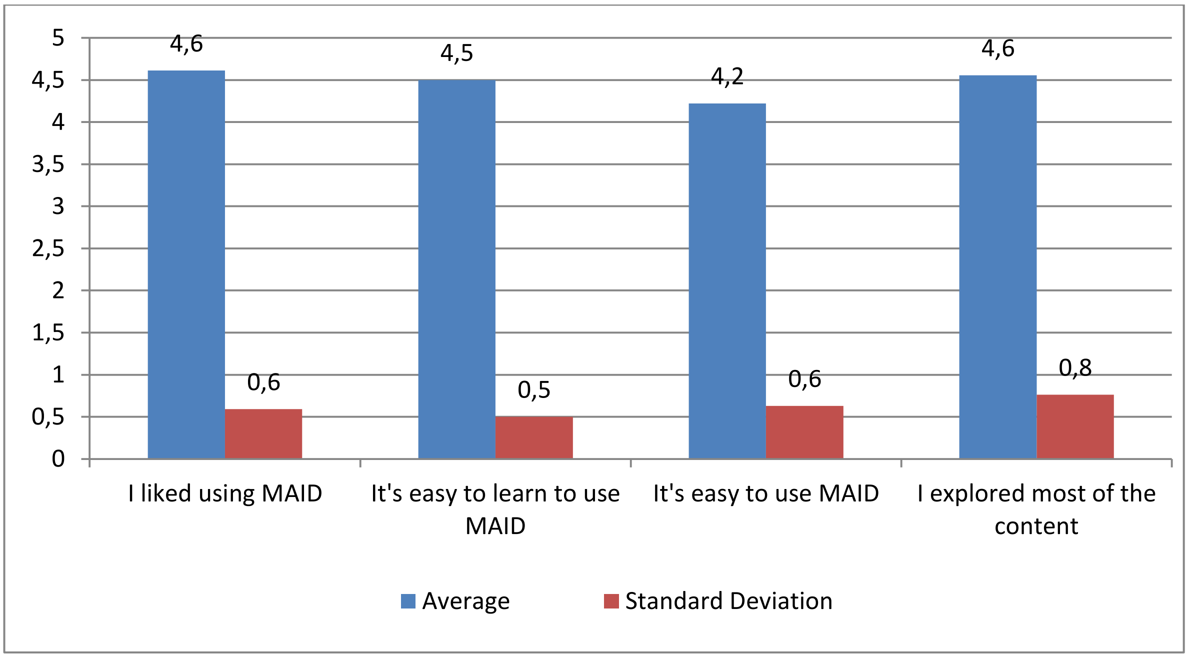


Chart 5.4 - General feedback

In the fifth question group, in the first section (Chart 5.5) the results were also highly positive, but not as consistent as in the previous question group. From the results, we can conclude that hand detection was not an obstacle for the users, the information and feedback supplied by MAID was simple and effective enough so that the users did not need much help to understand the interface. They also seem to grasp the relation between the gestures executed while using MAID and the corresponding real world action, resulting in a more engaging gameplay. This is an important element of the persuasive experience, so the results show that the majority of the users understood that the gestures are more than a simple interaction methodology, but a representation of simple daily actions.

One of the statements with the lowest average score regards the game terrain exploration, but it reached a positive result nonetheless, with a standard deviation of about one point. The terrain exploration process does not draw a fixed script of tasks to follow, depending on the users curiosity to find the intractable areas, in order to optimize the power consumption of the virtual home. Also the user is presented only with a needle gage (Figure 4.7) to evaluate how far he is from completing all the possible tasks, a state only reached when the needle reaches the far left side of the gage, on the light green area. From the results presented were positive, indicating that this setting is stimulating for most of the users.

The statement that had the lowest score and highest standard deviation regards the users eagerness to share their MAID interaction experience on Facebook, although the average result was moderately positive. From the oral feedback that we got from the users, they did not seem very eager to have a third party application like the MAID publishing content on their Facebook profile directly. We anticipated this response from the users, so the Facebook interaction interface that was presented in future studies did not publish the results directly to the users Facebook page, but publishes the content on the DEAP project gallery, giving the user the option to claim their accomplishments for their own profile.

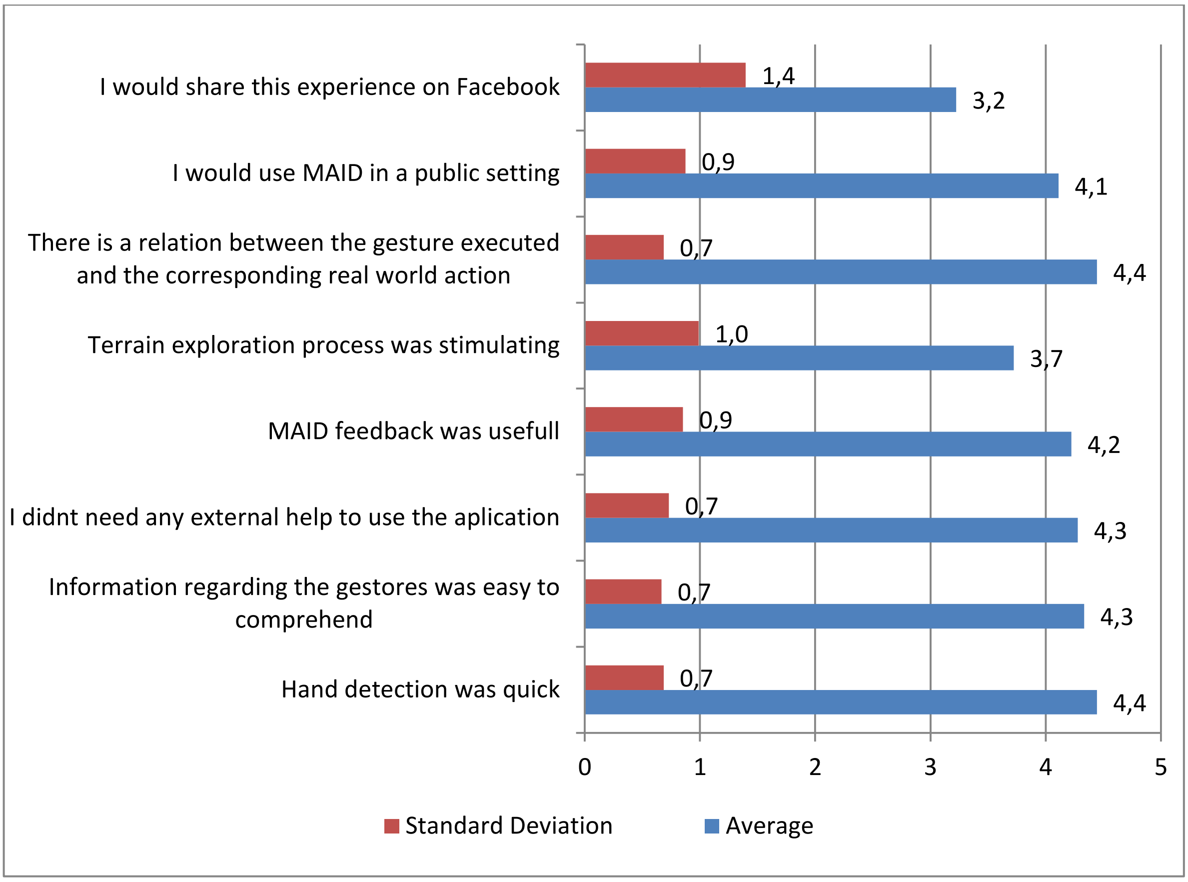


Chart 5.5 - Usability feedback

For the second part of the fifth question group (Chart 5.6) the average scores were high and consistent, so we can conclude that users mostly found that the gestures were overall easy to reproduce and also natural and appropriate regarding their context. The gesture that users seemed to struggle more to reproduce, and found less natural was the rotation gesture associated with the task of changing light bulbs. This particular sub-class of the main rotation gesture has the added complexity of having two phases, first the user has to perform a counter-clockwise gesture to unscrew the old light bulb, and then a clockwise gesture to screw in the new light bulb. Extra difficulties performing these actions were expected.

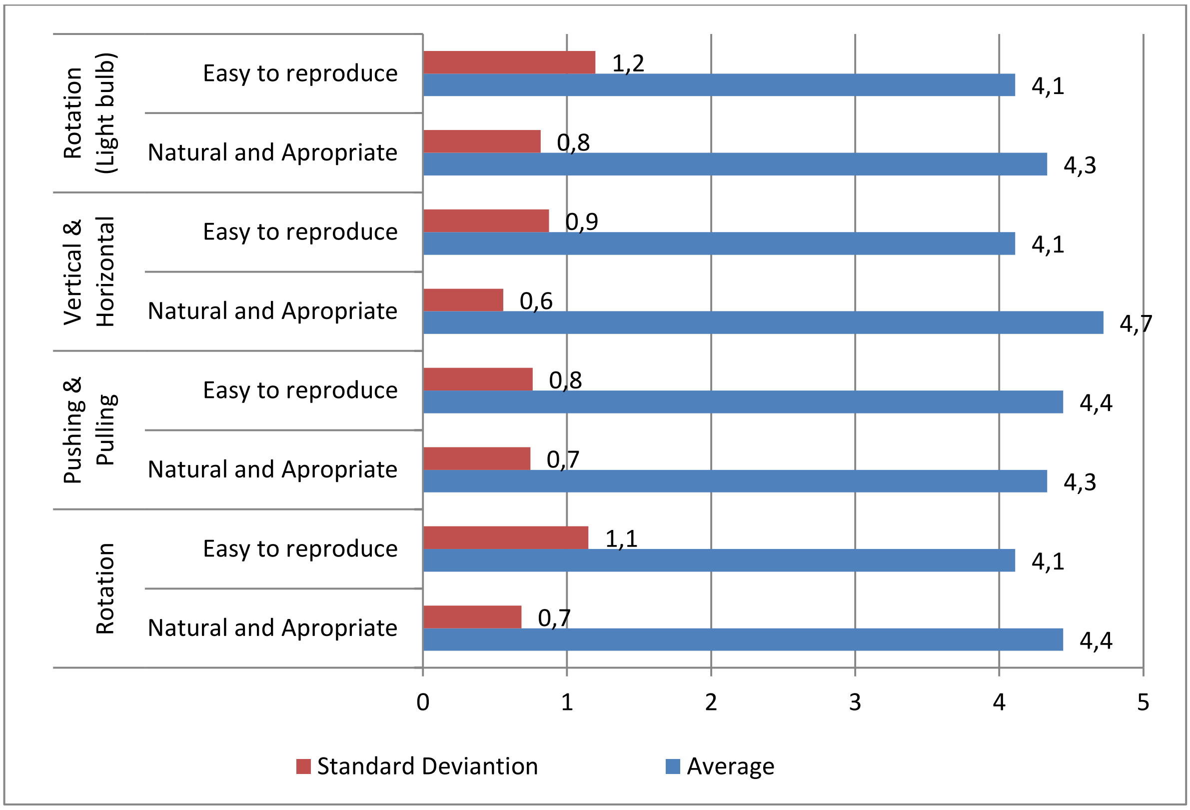


Chart 5.6 - Usability feedback regarding gesture types

In the sixth part of the questionnaire, the responses varied for most of the questions, but positive nonetheless. For the question “Did you obtain new knowledge with the given data?” most of the users answered “Yes” (Chart 5.7). The question “Did the user experience alert you to the consequences of the usage patterns?” had a very positive (Chart 5.8). From the answers to these two questions we can conclude that MAID was able to alert the majority of the user about the consequences their actions in this context. To the question “Did you already implement any of the suggestions?”, 66,7% of the users reported that they already implement most of the suggestions, two users already implemented all of the given suggestions and only one user reported implementing none of them (Chart 5.9). The last question of this question group was “Did MAID motivate you to implement given suggestions?”. Most of the responses to this question were divided between the “Yes” and “Maybe” answers, with an equal result of 37,5% and 25% of the users answered “No” (Chart 5.10). Although the number of “No” answers seem high, this can be justified by the fact that more than half the users already implement most of the suggestions, and maybe the ones that they do not implement just do not apply to their home.

The general results of this group of question indicate that MAID was able to influence the opinion of the users and serve the purpose of motivational tool.

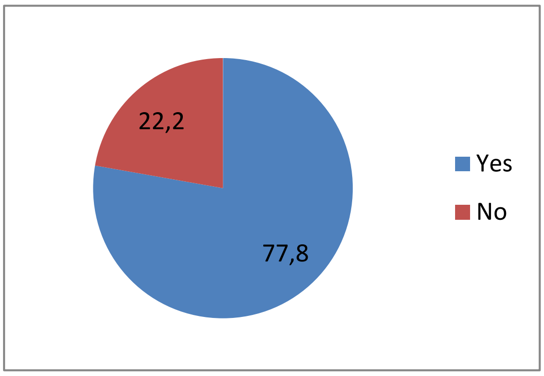


Chart 5.7 - Did you obtain new knowledge with the given data?

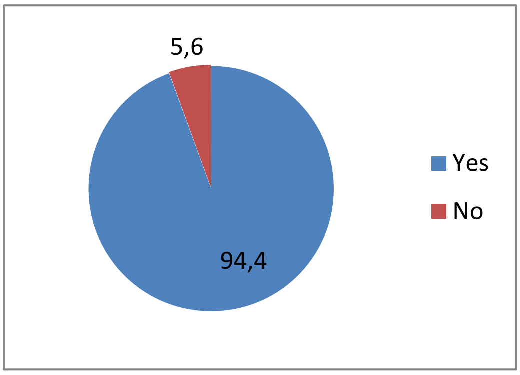


Chart 5.8 - Did the user experience alert you to the consequences of the usage patterns?

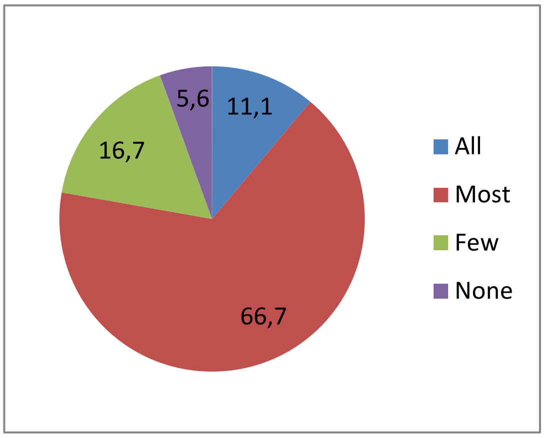


Chart 5.9 – Did you already implement any of the suggestions?

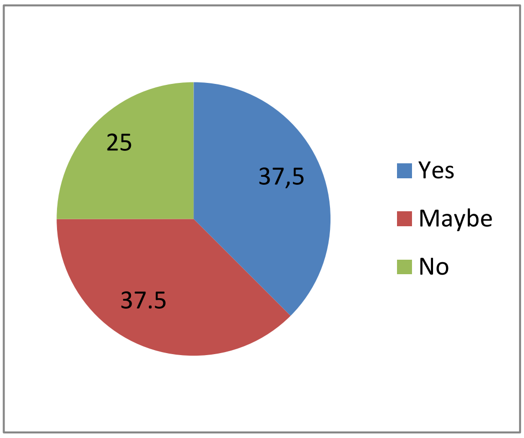


Chart 5.10 - Did MAID motivate you to implement given suggestions?

In the seventh question group, the users were asked to select a set of words that described their reaction to MAID. The results were very positive, with all of the positive words being select at least once and only two negative words being selected (Chart 5.11). The three most select words of the set were Useful, Innovative and Motivating. This is the result that would be expected for a persuasive interface. The two negative words selected were "Boring" and "Irritating". According to our observations and feedback from the users, this is associated with a detail of the implementation of this version of the prototype where a user that just the direct interaction stage (see 4.2.8) is presented with a text information screen that he is forced to contemplate for too long. Furthermore, during this period if the user lowered is hand in a way that lead to the loss of tracking, the system would present the calibration interface, so this forced the user to keep his tracked hand up during this period. This flaw was corrected on the next version of the application for the next user study, by cutting the text presentation time down to the minimum necessary for the user to read it. This was done by tweaking the configurations of the application. Also the user is no longer asked to recalibrate the hand tracking during this period.

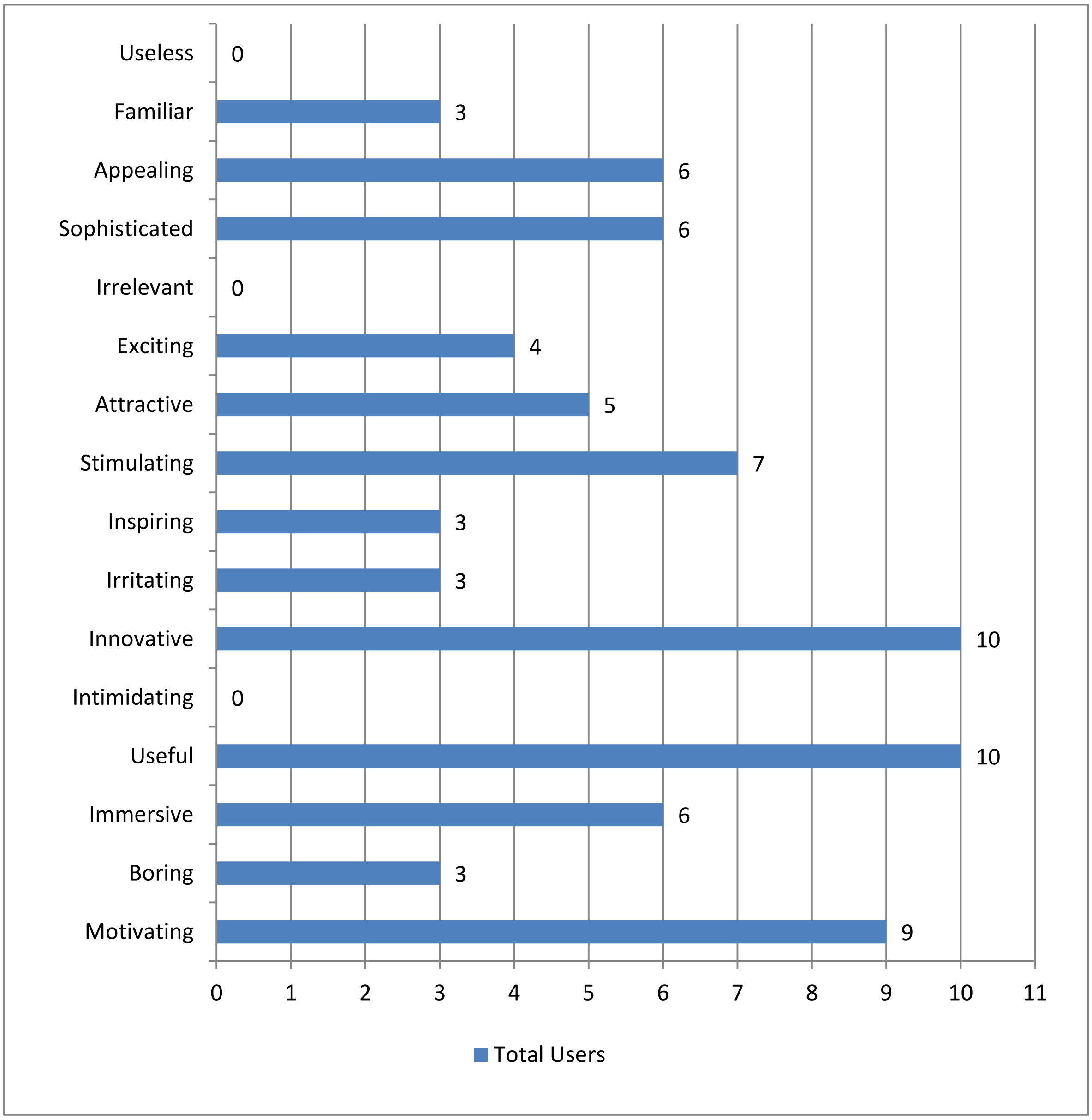


Chart 5.11 - Reaction to MAID

In the eighth question group, the users were asked to select a set of words that described their interaction with MAID and the results were positive (Chart 5.12), with only one positive word, "Comfortable" not being selected at all. Three negative words were selected: "Frustrating", "Uncomfortable" were selected by two users and "Confusing" by one user. Since the word "long" was not selected, it seems that the discomfort is not caused by a lengthy interaction. Another possible explanation for the selection of these words was the previously stated design flaw. As for the selection of the word "Confusing" by one user, the impact is mitigated by the popular choice of the phrase "easy-to-use". Other popular word choices were: "Simple", "Intuitive" and "Fun".

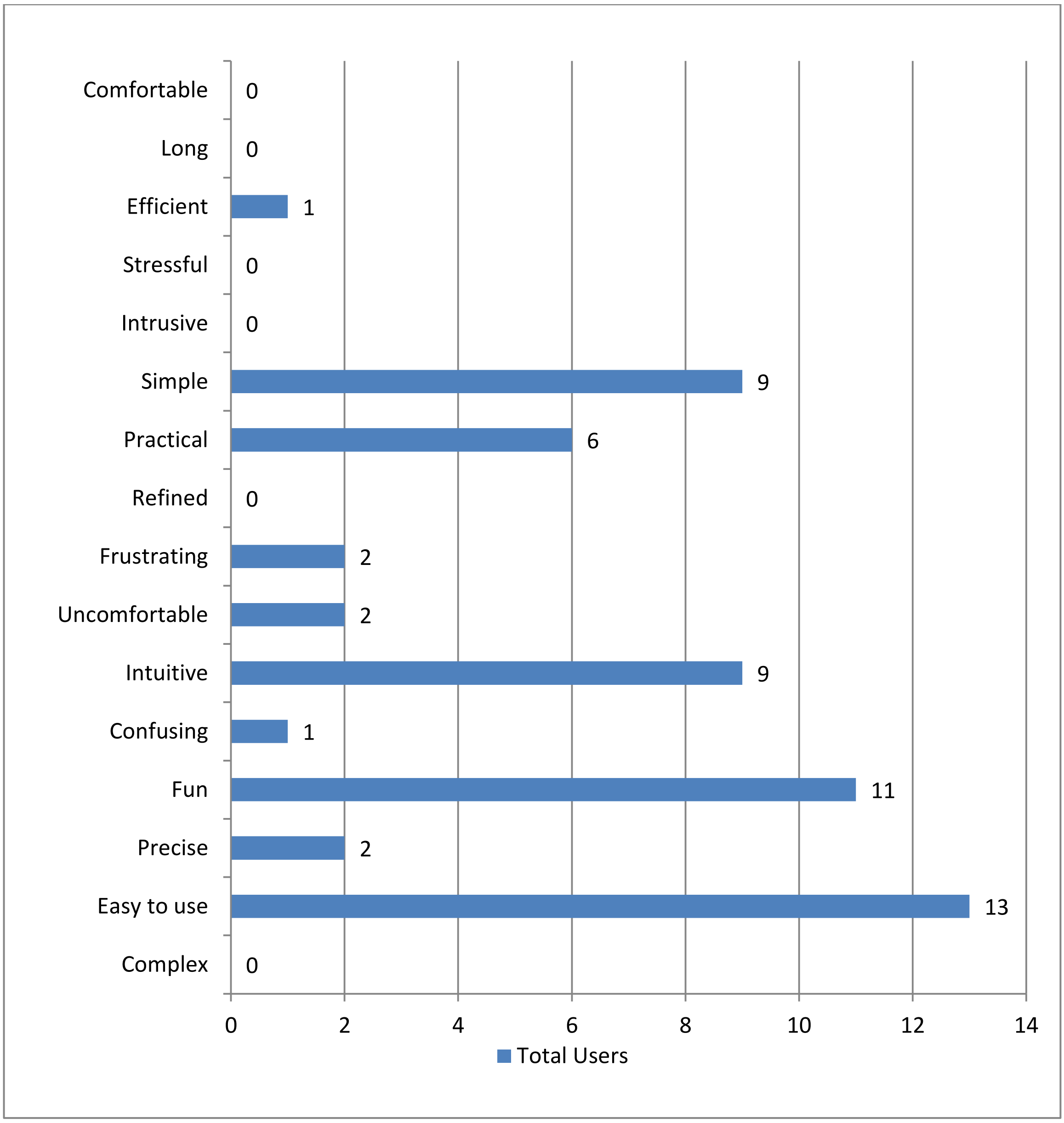


Chart 5.12 - MAID usage

The results from this user study were positive, as we can observe in the presented results. The users were eager to explore the contents presented in MAID using the gesture based interface and at the end of the interaction most of them seemed to be persuaded to implement the rehearsed actions at their home.

One of questions that raised more confusion during this test was on the second part of the fifth question group, regarding the use of the gesture. Some participants did not understand the difference between the basic rotation gesture and the rotation gesture used for the light bulbs. In the next user study there was no differentiation between these rotation gestures.

For this event, the Facebook interface was not yet available, but it was integrated in system for the next user study.

5.2. Expo FCT User Study

A user evaluation study was performed during an open day event, attended by 6000 high school students, at University Campus of FCT. Twenty six users aged 15-28 (average of 17.5) participated in the evaluation procedures (21 male and 5 female). All of them were familiarized with new technologies. This set of users is younger and larger in number when compared to the first user study.

For this evaluation study we had enough space and control over lighting conditions, which allowed us to use a projector and get a projection area large enough so that the audience does not need to crowd the user interacting with the MAID. The projector was overhead mounted to maximize the unobstructed area needed for the projection. The Kinect was mounted below the projection screen, and the users were positioned at a minimum distance of 2 meters from the sensor, with about 1.8m radius with no obstruction, allowing freedom of movements. Due to space constraints (other activities were occurring in the same room), audience was kept on the right side of the MAID deploying area.

The test sessions were conducted by two researchers, who played the roles of facilitator and observer. The first one had a more active role, giving an initial briefing and instructions to the participants and providing assistance for any problems that users might face. The second researcher focused on observing the way the tests unfolded, and how users reacted and interacted with the system.

Before starting to use the application, users were informed about the objectives of the test. After a short description of the application and an explanation of the goals to be achieved, users were encouraged to explore MAID. They could do it for as long as they liked, as in a real setting.

At the end of each test, participants were asked to answer a questionnaire to assert their reactions, gather some personal information and collect their feedback and comments.

5.2.1. Questionnaire

The questionnaire used was very similar to the one used in the previous study, the main difference between them is that this one needed to be shorter and faster to fill in, due to the expected amount of users attending the event(see annex 2). The first part covered basic personal information like the participant's age, gender, familiarity with new technologies and frequency of use of Microsoft Kinect. The second group of questions was related with general feedback and usability issues, as well as with the appropriateness of the gestures. Users were asked to rate statements, using a five-point Likert-type scale, which ranged from strong disagreement (1) to strong agreement (5). The third part of the questionnaire included questions about the users' energy consumption attitudes and how the experience motivated them to change their behaviour. The fourth part of the questionnaire was based on the Microsoft "Product Reaction Cards", in order to capture the user's feelings towards the MAID. Users were asked to choose the words that best described their experience while using the application from a list of 20 words, a set that included about 60% of words considered positive and 40% considered negative. The participants could choose any number of words that they deemed adequate to describe their reaction. Finally, the questionnaire also included an open question, in order to gather comments and recommendations regarding future developments of additional features and to obtain a more general evaluation of the system. Finally, the questionnaire also included an open question, in order to gather comments and recommendations regarding future developments of additional features and to obtain a more general evaluation of the system.

5.2.2. Results and Discussion

Many participants had never used the MS Kinect before (58%), as shown in Chart 5.13, and MAID had taken on the rule of introducing them to gesture based interfaces. Contrasting with the first user study, where every user had already at least tried the sensor once.

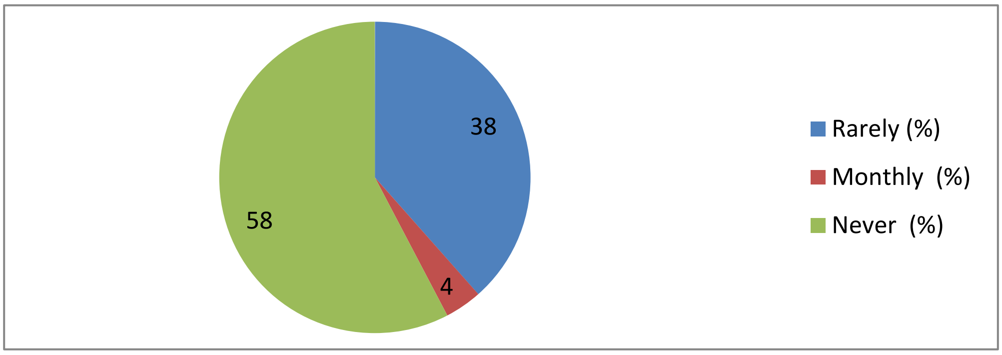


Chart 5.13 - Kinect usage

Regarding the second group of questions, the results were very positive, as it can be observed in Chart 5.14 and Chart 5.15. The majority of participants liked to use MAID and found it easy to learn and to use. They considered the feedback and the information provided by the application useful and they were willing to use MAID in public spaces. Observations revealed that users had fun while using MAID and their friends, in the audience, were commenting and giving suggestions. These results corroborate the finding of the first user study.

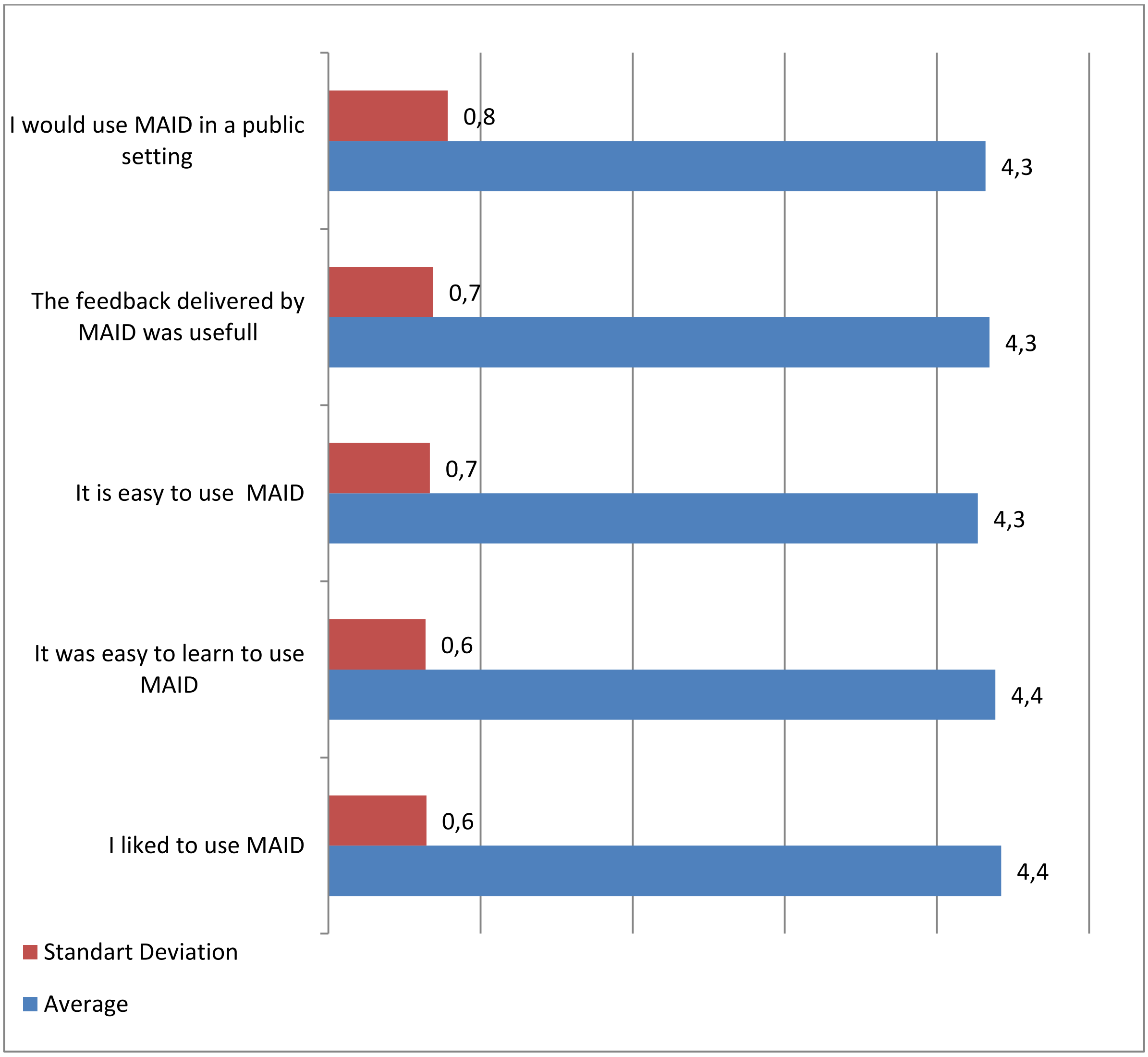


Chart 5.14 - Summary of results from general feedback and usability.

Results regarding the gesture commands were also very positive (Chart 5.15). Participants considered the gestures natural and appropriate to mimic the real action, as well as easy to perform. Horizontal sweep was the gesture that marginally achieved the most positive results. These results are generally very similar to the ones on the first user study, but since some of the questions are different, a direct comparison of the results between question cannot be established. The main difference between the two user test, regarding this question group, is that in this one we evaluated vertical and horizontal gestures independently and do not evaluate the Rotation gesture sub-class used for the light bulb objects. The results were slightly more consistent than the ones in first user study, with all the gestures having very similar results, this could also be caused by the number of the users in this study being greater.

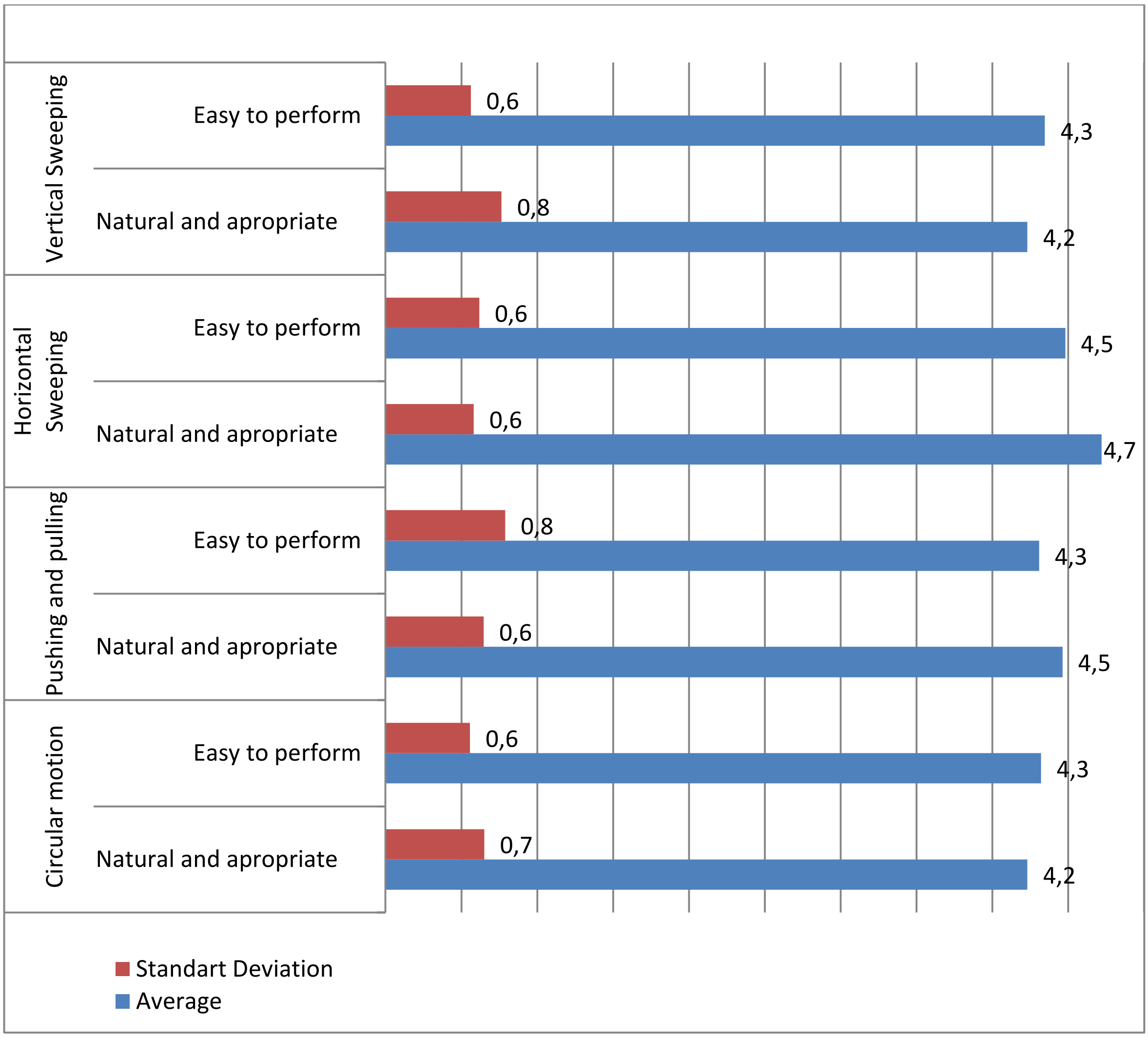


Chart 5.15 - Summary of results regarding gestures commands.

The questionnaire also included questions about the users' energy consumption attitudes and how the experience motivated them to change their behaviors. All participants except two considered that MAID provided them with new useful information and knowledge regarding domestic energy consumption and all participants, except three, answered that their experience using MAID made them aware of the consequences of their appliance usage patterns. Half of the participants were already implement most of the MAID suggestions in their everyday lives before the user tests, and 27% of them implemented some of the suggestions. Compared with the previous study, there are less participants reporting that they implement most of the suggestions, this is probably do to the lower average age of the users of this study. But most importantly, 50% of the participants reported that they were motivated by MAID to implement the given suggestions, and 46% answered "maybe". These results show a difference in the number of users that answered "no" (11%) when compared to the previous user study (26,7%), so the younger participants seem more eager to implement the given suggestions.

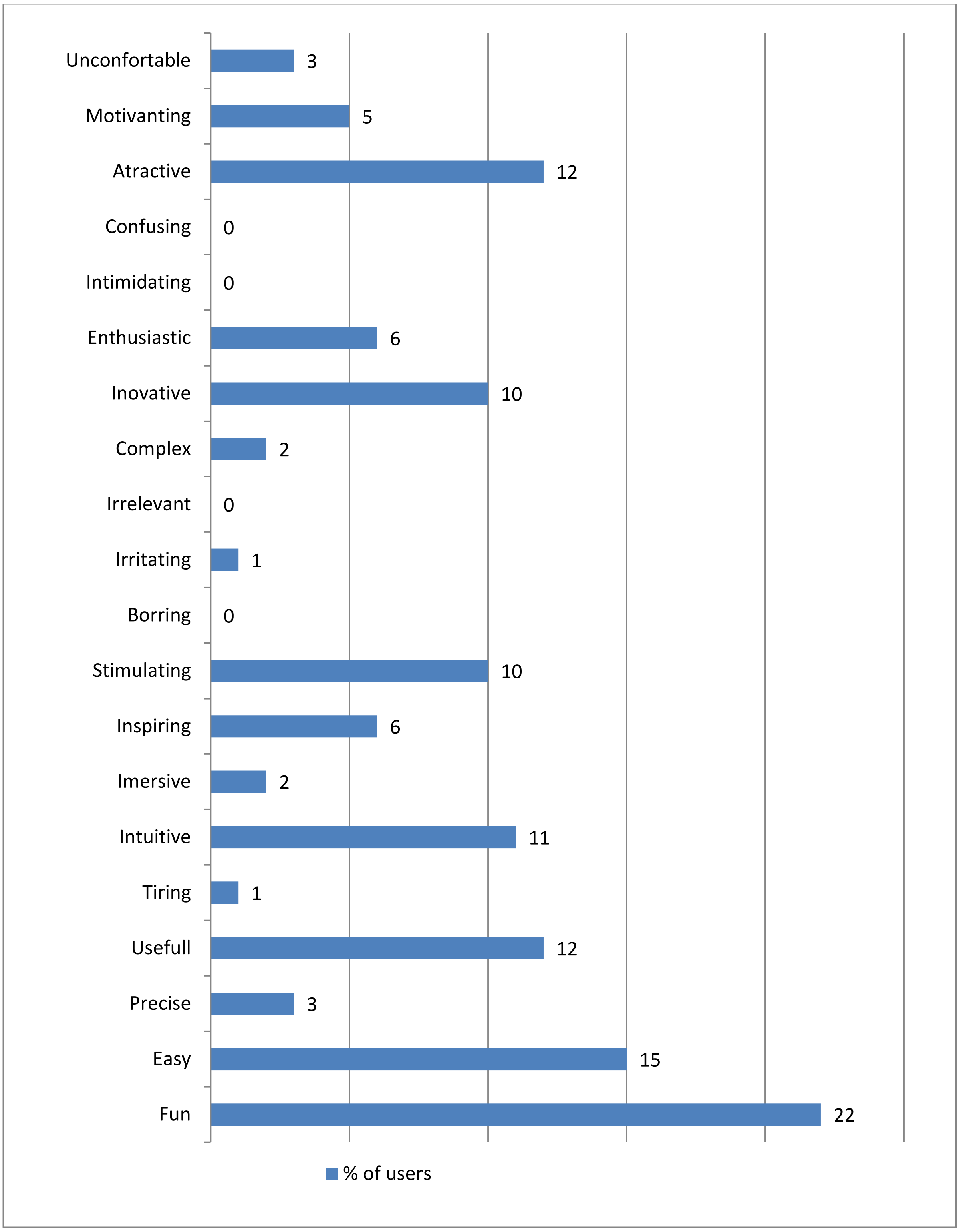


Chart 5.16 - User experience feedback

From the analysis of question based on the Microsoft "Product Reaction Cards" we concluded that all participants held positive feelings when classifying their experience using the MAID (Chart 5.16), 22 users selected the word "Fun", which is an important quality of an application of this kind to be deployed in public spaces. Other positive top selected words were: Easy (15), Attractive (12) and Useful (12). All positive words included in the set were selected by the users, and given the word choice, the users valued the overall look and usability aspects of the application. Even though most of the feedback was positive, some negative words were also selected by only a few participants: Uncomfortable (3), Complex (2), Tiring (1) and Irritating (1).

A few users reported that the hand gesture interface became tiring after a few minutes of interaction. Applications like MAID should not be used for a long time and users must be able to accomplish the game objectives in a few minutes. Although the word “Boring” was not selected this time, words like “Uncomfortable” and “Irritating” still appear in the result, despite the correction and adjustments done since the previous study. One of the reasons that could explain the selection of the negative words is the average age of the users, being more impatient and easily distracted. Another reason would be that many participants never used the connect (58%).

The Facebook interface was also well received by the users, but only two participants took individual pictures. Most participants took group pictures with their friends. A total of fourteen photos were published, but most participants seem to forget to later visit the project Facebook page to check their pictures, since not many photos were tagged or liked. After the shot was taken, the users were informed of where the picture would be available, but the only few that interacted with their pictures on Facebook were the ones that immediately visited the page, either using their smart phones, or by using a laptop borrowed. This behaviour limits the dissemination of the information through the social network graph, future work is needed to reach a solution for this issue.

The participants showed great enthusiasm in exploring the gameplay, during this study. The reaction of the users was very positive, with the MAID installation constantly gathering an audience.

6. Conclusions and Future Work

This Master Thesis resulted in the creation of MAID, an interactive public ambient display system, driven to motivate behaviour changes regarding domestic energy consumption, through a persuasive game interface based on gesture recognition technology. The main objective of MAID is to instruct (or remind) users about simple procedures to save energy, showing them how easy it is to have a huge impact on the environment by taking simple actions in their daily lives.

The MAID implementation is based on a modular and highly configurable public ambient display framework. The flexibility of the architecture can even allow it to be repurposed to target different interaction studies, with themes and goals beyond the scope of this thesis.

In order to reach the development of MAID, a series of small prototypes were developed to test possible technological solutions. This was the case of KPong, a small scale test application developed to test the potential of the Microsoft Kinect.

The developed MAID system was evaluated in two separate user studies, first a smaller scale test and later a larger test during the open day at our University Campus, attended by 6000 high-school students. Results were very positive in both studies. Most participants in the user tests liked to use MAID found it fun and easy to use and considered that MAID provided them with new useful information regarding domestic energy consumption, motivating them to implement the suggestions provided by MAID in their everyday lives. Therefore, MAID seems to have the potential to influence people to change their behaviour. However, we are aware that further tests need to be conducted in order to measure the persuasive effect of the game in the long term, to evaluate if the behaviour change persisted.

According to the evaluation results, the designed gesture commands seem to be appropriate for the purpose and easy to perform. Users did not have difficulties understanding the relation between the interface gesture and the corresponding action in the real world.

MAID was demonstrated, as a creative showcase, at the 8th International conference on Advanced in Computer Entertainment Technology (ACE), originating a publication on ACM DL (Salvador & Romão, 2011). A full paper will also be presented at the ACE 2012 and published by Springer (Salvador, Romão, & Centieiro, 2012) (see Annex 3).

The Facebook interaction component of MAID would benefit from some improvements, like giving the user the possibility to confirm or retake a shot, before it is sent to Facebook, or even cancel the whole process without taking any shot. This would increase the usability of this functionality. It also needs a mechanism that would allow the user to link the photograph taken directly to his account. This could be done with a QR code, were the MAID after taking the shot and sending it to Facebook, would generate a unique QR CODE that would link to the picture, so that the user could use a handheld device, such as his mobile phone to open the link and see the picture through his Facebook account.

Conduct user studies that would evaluate the effectiveness of the persuasion methodology applied in the MAID in the long term, evaluating the persistency potential behaviour change.

The need for a calibration routine in MAID could be eliminated be migrating the application to a official Microsoft SDK framework based implementation, it would also allow further additions and improvements that would require more precise and reliable recognition. Without the need for a calibration routine for skeleton tracking, it could be a viable tool for new forms of interaction. When the development of MAID started the official SDK had not yet been released, and when it was finally released the migration effort would be too costly at the time, since the open-source based MAID implementation would have to suffer a dramatic rewriting of most of the source code.

Bibliography

Global Warming Articles News and Facts. (2011). Retrieved from Global Warming: http://www.globalwarmingarticle.com/globalwarmingnews/

Human–computer interaction. (2011). Retrieved from Wikipedia: http://en.wikipedia.org/wiki/Human-computer_interaction

Kinect. (2011). Retrieved from Wikipedia, the free encyclopedia: http://en.wikipedia.org/wiki/Kinect

Adam, D. (2007, February 19). *Climate change: scientists warn it may be too late to save the ice caps.* Retrieved from The Guardian: http://www.guardian.co.uk/environment/2007/feb/19/frontpagenews.climatechange

Bang, M., Gustafsson, A., & Katzeff, C. (2007). Promoting new patterns in household energy consumption with pervasive learning games. (pp. 55--63). Springer-Verlag.

Benedek, J., & Miner, T. (2002). Measuring Desirability: New methods for evaluating desirability in a usability lab setting, volume=2003, url=http://www.microsoft.com/usability/UEPostings/DesirabilityToolkit.doc, number=03/04/2003, journal=Proceedings of UPA Usability Professional Association. Microsoft Corporation, 1 Microsoft Way, Redmond, WA 98052.

Brignull, H., & Rogers, Y. (2003). Enticing people to interact with large public displays in public spaces. *INTERACT’03*, (pp. 17--24). Zurich, Switzerland.

de Oliveira, R., Cherubini, M., & Oliver, N. (2010). MoviPill: improving medication compliance for elders using a mobile persuasive social game. *UbiComp 2010* (pp. 251--260). Copenhagen, Demark: ACM.

Dillahunt, T., Becker, G., Mankoff, J., & Kraut, R. (2008). Motivating Environmentally Sustainable Behavior Changes with a Virtual Polar Bear. *Pervasive 2008 Workshop on Pervasive Persuasive Technology and Environmental Sustainability* , (pp. 58-62). Sydney,Australia.

Eisenhauer, M., Lorenz, A., & Zimmermann, A. (2008). Interaction-Kiosk For Open Human-Computer Interaction With Pervasive Services. *Third International Conference on Persuasive Technology*, (pp. 134-137). Oulu, Finland.

Fogg, B. (2009) (A). A behavior model for persuasive design. *Persuasive '09* (pp. 40:1--40:7). ACM.

Fogg, B. (2009)(B). The Behavior Grid: 35 ways behavior can change. *Persuasive '09* (pp. 42:1--42:5). ACM.

Fogg, B. J. (2002, December). Persuasive technology: using computers to change what we think and do. *Ubiquity, 2002*(December).

Fogg, B. J. (2008) (A). Mass Interpersonal Persuasion: An Early View of a New Phenomenon. *PERSUASIVE '08* (pp. 23--34). Berlin, Heidelberg: Springer-Verlag.

Fogg, B. J. (2009). Creating persuasive technologies: an eight-step design process. *Persuasive '09* (pp. 44:1--44:6). New York, USA: ACM.

Froehlich, J., Dillahunt, T., Klasnja, P., Mankoff, J., Consolvo, S., Harrison, B., et al. (2009). UbiGreen: investigating a mobile tool for tracking and supporting green transportation habits. *CHI '09* (pp. 1043--1052). Boston, USA: ACM.

Gustafsson, A., BÃ¥ng, M., & Svahn, M. (2009). Power explorer: a casual game style for encouraging long term behavior change among teenagers. *ACE '09* (pp. 182--189). Athens, Greece: ACM.

Hardy, J., Rukzio, E., & Davies, N. (2011). Real world responses to interactive gesture based public displays. *MUM '11* (pp. 33--39). Beijing, China: ACM.

IJsselsteijn, W., de Kort, Y., Midden, C., Eggen, B., & van den Hoven, E. (2006). Persuasive Technology for Human Well-Being: Setting the Scene. *PERSUASIVE'06. 3962*, pp. 1-5. Eindhoven, The Netherlands: Springer Berlin / Heidelberg.

Jacucci, G., Morrison, A., Richard, G. T., Kleimola, J., Peltonen, P., Parisi, L., et al. (2010). Worlds of information: designing for engagement at a public multi-touch display. *CHI '12* (pp. 2267--2276). Atlanta, Georgia, USA: ACM.

Jahn, M. J.-A. (2010). The Energy Aware Smart Home. *FutureTech 2010.* NY: IEEE Press.

Kuznetsov, S., & Paulos, E. (2010). UpStream: motivating water conservation with low-cost water flow sensing and persuasive displays. *CHI '11* (pp. 1851--1860). Atlanta, Georgia, USA: ACM.

Laitner, J. E.-M. (2009). Examining the Scale of the Behaviour Energy Efficiency Continuum. *European Council for an Energy Efficient Economy Con-ference*, (pp. 217-223). Cote d'Azur, France.

Loke, S. W., Singh, J., & Le, H. (2008). Context-Aware Pervasive Persuasive Systems For Managing Water and Energy Usage, and CO2 Emissions: Multi-Leveled Policies, Goals and Expert Systems Shell Approach. *Pervasive '08.* Sydney, Australia.

Mattern, F., Staake, T., & Weiss, M. (2010). ICT for green: how computers can help us to conserve energy. *e-Energy '10* (pp. 1--10). Passau, Germany: ACM.

Microsoft. (2012). *Kinect for XBox 360*. Retrieved Jully 2012, from http://www.xbox.com/pt-PT/kinect

Midden, C. J., Kaiser, F. G., & Teddy McCalley, L. (2007,). Technologys Four Roles in Understanding Individuals Conservation of Natural Resources. *Journal of Social Issues, 63*, 155--174.

Midden, C. J., McCalley, L. T., Ham, J. R., & Zaalberg, R. (2008). Using persuasive technology to encourage sustainable behavior. *Prevasive '08.* Aberdeen: Conference Organization.

Miller, T. M., Rich, P., & Davis, J. (2009). ADAPT: audience design of ambient persuasive technology. *CHI EA '09* (pp. 4165--4170). Boston, USA: ACM.

Myers, B. A. (1998, March). A brief history of human-computer interaction technology. *interactions, 5*(2), 44--54.

Penny, H., & Underwood, D. (2008). Understanding motivation and enabling action. *Pervasive '08,* (pp. 1-4). Sydney, Australia.

Reeves, B., Cummings, J. J., Scarborough, J. K., Flora, J., & Anderson, D. (2012). Leveraging the engagement of games to change energy behavior. *CHI '11* (pp. 354-358). Vancouver, Canada: ACM.

Reitberger, W., Ploderer, B., Obermair, C., & Tscheligi, M. (2007). The percues framework and its application for sustainable mobility. *PERSUASIVE'07* (pp. 92--95). Palo Alto, USA: Springer-Verlag.

Revelle, G., Reardon, E., Green, M. M., Betancourt, J., & Kotler, J. (2007). The use of mobile phones to support childrens literacy learning. *PERSUASIVE'07* (pp. 253--258). Palo Alto, USA: Springer-Verlag.

Rubegni, E., Memarovic, N., & Langheinrich, M. (2011). Talking to Strangers: Using Large Public Displays to Facilitate Social Interaction. *HCII 2011.* Springer.

Salvador, R., & Romão, T. (2011). Lets move and save some energy. *ACE '11* (pp. 86:1--86:2). Lisbon, Portugal: ACM.

Salvador, R., Romão, T., & Centieiro, P. (2012). A Gesture Interface Game for Energy Consumption Awareness. *ACE '12* (p. Article Nº.78). Kathmandu, Nepal: Springer.

Shoemaker, G., Tsukitani, T., Kitamura, Y., & Booth, K. S. (2010). Whole body large wall display interfaces. *CHI EA '10* (pp. 4809--4812). Atlanta, USA: ACM.

Spaargaren, G., & Van Vliet, B. (2000). Lifestyles, consumption and the environment: The ecological modernization of domestic consumption. *Environmental Politics, 9,* 50-76.

Stein, J. (2008). TerraPed: A Design Concept for Sustainability. *Pervasive '08.* Sydney, Australia.

Sutaria, R., & Deshmukh, A. (2008). Taking the Guesswork out of Environmentally. *Pervasive '08.* Sydney, Australia.

Takayama, C., Lehdonvirta, V., Shiraishi, M., Washio, Y., Kimura, H., & Nakajima, T. (2009). ECOISLAND: A System for Persuading Users to Reduce CO2 Emissions. *Future Dependable Distributed Systems, Software Technologies for, 0,* 59-63.

Vatavu, R.-D. (2012, 01 01). Nomadic gestures: A technique for reusing gesture commands for frequent ambient interactions. *Journal of Ambient Intelligence and Smart Environments, 4,* 79--93.

Vogel, D., & Balakrishnan, R. (2004). Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. *UIST '04* (pp. 137--146). Santa Fe, USA: ACM.

Annex 1 – MAID Questionnaire - First User Study

Motion-Based Ambient Interactive Display (MAID)

Questionário de Avaliação

Todos os dados recolhidos são confidenciais e não serão utilizados para qualquer outra finalidade.

1. Dados Pessoais

1.1. Idade:_______

1.2. Sexo:

☐ Masculino ☐ Feminino

2. Grau de Instrução

☐ Doutoramento/Pós-Doutoramento	☐ Mestrado/Pós-Graduação	☐ Licenciatura
☐ Ensino Secundário	☐ 3º Ciclo do Ensino Básico	☐ 2º Ciclo do Ensino Básico
☐ 1º Ciclo do Ensino Básico		

3. Novas Tecnologias

3.1. Que tipo de tecnologias utiliza?

☐Computador ☐Diariamente ☐Semanalmente ☐Mensalmente ☐Raramente	☐Smart Phones ☐Diariamente ☐Semanalmente ☐Mensalmente ☐Raramente
☐Consolas de jogos ☐Diariamente ☐Semanalmente ☐Mensalmente ☐Raramente	☐Internet ☐Diariamente ☐Semanalmente ☐Mensalmente ☐Raramente
☐Outra. Qual? ________________________________ ☐Diariamente ☐Semanalmente ☐Mensalmente ☐Raramente	

3.2. É a primeira vez que interage com a Kinect?

☐ Sim ☐ Não

3.2.1 Se respondeu não, com que frequência utiliza a Kinect?

☐Diariamente ☐Semanalmente ☐Esporadicamente ☐Experimentei

Motion-Based Ambient Interactive Display (MAID)
Questionário de Avaliação

Responda às seguintes questões fazendo um círculo em volta do número que melhor representa a sua opinião acerca da aplicação que acaba de experimentar.

4. Aspectos Gerais					
	Discordo				Concordo
4.1. Gostei de utilizar o MAID.	1	2	3	4	5
4.2. É fácil aprender a usar o MAID.	1	2	3	4	5
4.3. É fácil utilizar o MAID.	1	2	3	4	5
4.4. Explorei a maioria dos conteúdos do MAID.	1	2	3	4	5

5. Usabilidade					
	Discordo				Concordo
5.1. A calibração da detecção das mãos foi rápida.	1	2	3	4	5
5.2. A informação relativa aos gestos a executar é de fácil compreensão.	1	2	3	4	5
5.3. Consigo usar a aplicação sem recorrer a ajuda exterior.	1	2	3	4	5
5.4. O feedback fornecido pela aplicação foi útil.	1	2	3	4	5
5.6. O processo de exploração do cenário apresentado é estimulante.	1	2	3	4	5
5.7. Existe relação entre a simplicidade de cada um dos gestos executados no MAID e a correspondente acção a executar no ambiente real.	1	2	3	4	5
5.8. Utilizaria o MAID num ambiente público	1	2	3	4	5
5.9. Partilharia esta experiência no Facebook, se tal funcionalidade estivesse disponível.	1	2	3	4	5

5.10. O seguinte conjunto de perguntas é dedicado aos gestos utilizados no MAID:

		Discordo			Concordo	
1 Gestos de rotação, associados a objectos como o botão do forno ou o termóstato	São naturais e apropriados, dada a acção a executar.	1	2	3	4	5
	São de fácil reprodução	1	2	3	4	5
2 Gestos com movimentos em profundidade , associados a objectos como interruptores ou o transformador.	São naturais e apropriados, dada a acção a executar.	1	2	3	4	5
	São de fácil reprodução	1	2	3	4	5
3 Gestos com movimentos em verticais ou horizontais , associados a objectos como a janela ou cortinas:	São naturais e apropriados, dada a acção a executar.	1	2	3	4	5
	São de fácil reprodução	1	2	3	4	5
4 Gestos associados à mudança das lâmpadas:	São naturais e apropriados, dada a acção a executar.	1	2	3	4	5
	São de fácil reprodução	1	2	3	4	5

Motion-Based Ambient Interactive Display (MAID)
Questionário de Avaliação

6. Poupança de energia no contexto doméstico

6.1. Aprendeu com a informação disponibilizada pelo MAID?

☐ Sim ☐ Não

6.2. A experiência de utilização do MAID sensibiliza para as consequências de certos padrões de utilização de equipamentos?

☐ Sim ☐ Não

6.3. Já implementava as sugestões dadas pelo MAID?

☐ Todas ☐ A maioria ☐ Poucas ☐ Nenhuma

6.3.1. Caso não tenha respondido "Todas", acha que o MAID o motivará a implementar estas sugestões no seu dia-a-dia?

☐ Sim ☐ Talvez ☐ Não

7.Reacção ao MAID

Escolha as expressões que melhor definem o MAID.

☐ Motivante	☐ Aborrecida	☐ Imersiva	☐ Útil
☐ Intimidante	☐ Inovadora	☐ Irritante	☐ Inspiradora
☐ Estimulante	☐ Atractiva	☐ Entusiasmante	☐ Irrelevante
☐ Sofisticada	☐ Apelativo	☐ Familiar	☐ Inútil

8.Utilização do MAID

Escolha as expressões que melhor definem a sua experiência na utilização do MAID.

☐ Complexa	☐ Fácil de usar	☐ Precisa	☐ Divertida
☐ Confusa	☐ Intuitiva	☐ Desconfortá-vel	☐ Frustrante
☐ Refinada	☐ Prática	☐ Acessível	☐ Intrusiva
☐ Stressante	☐ Eficaz	☐ Grande	☐ Confortável

9. Sugestões e Comentários

____

____

____

Muito obrigado pela sua colaboração.

Annex 2 – MAID Questionnaire - Expo FCT user study

Motion-Based Ambient Interactive Display

Questionário de Avaliação

Todos os dados recolhidos são confidenciais e não serão utilizados para qualquer outra finalidade.

1. Dados Pessoais e Novas Tecnologias

1.1. Idade: _______ 1.2. Sexo: ❑ Masculino ❑ Feminino

1.3. Estás familiarizado com as novas tecnologias? ❑ Sim ❑ Não

1.4. É a primeira vez que interages com a Kinect? ❑ Sim ❑ Não

1.4.1. Caso tenhas respondido que **não** com que frequência utilizas a Kinect?

❑ Diariamente ❑ Semanalmente ❑ Mensalmente ❑ Raramente

Avalia as seguintes afirmações fazendo **um círculo em volta do número** que melhor representa a tua opinião acerca da aplicação que acabaste de experimentar.

2. Usabilidade

		Discordo				Concordo
2.1. Gostei de utilizar o MAID.		1	2	3	4	5
2.2. É fácil aprender a usar o MAID.		1	2	3	4	5
2.3. É fácil usar o MAID.		1	2	3	4	5
2.4. O *feedback* fornecido pela aplicação foi útil.		1	2	3	4	5
2.5. Utilizaria o MAID num local público.		1	2	3	4	5
2.6. O seguinte conjunto de perguntas é dedicado aos gestos utilizados no MAID.						
2.6.1. Gestos de rotação, associados a objectos como o botão do forno ou o termóstato:	É natural e apropriado à acção a executar.	1	2	3	4	5
	São de fácil reprodução.	1	2	3	4	5
2.6.2. Gestos com movimentos em profundidade, associados a objectos como interruptores ou transformadores:	É natural e apropriado à acção a executar.	1	2	3	4	5
	São de fácil reprodução.	1	2	3	4	5
2.6.3. Gestos com movimentos verticais ou horizontais, associados a objectos como a janela ou cortinas:	É natural e apropriado à acção a executar.	1	2	3	4	5
	São de fácil reprodução.	1	2	3	4	5
2.6.4. Gestos associados à mudança das lâmpadas:	É natural e apropriado à acção a executar.	1	2	3	4	5
	São de fácil reprodução.	1	2	3	4	5

3. Poupança de Energia no Contexto Doméstico

3.1. Aprendeste com a informação disponibilizada pelo MAID? ❑ Sim ❑ Não

3.2. A experiência de utilização do MAID sensibiliza-te para as consequências de utilização de equipamentos? ❑ Sim ❑ Não

3.3. Já fazias algumas das sugestões dadas pelo MAID? ❑ Todas ❑ A maioria ❑ Poucas ❑ Não

3.3.1. Caso não tenhas respondido "Todas", achas que o MAID te motivará a implementar estas sugestões no teu dia-a-dia? ❑ Sim ❑ Talvez ❑ Não

Motion-Based Ambient Interactive Display

Questionário de Avaliação

Todos os dados recolhidos são confidenciais e não serão utilizados para qualquer outra finalidade.

4. Reacção ao MAID

Escolhe as expressões que melhor definem a tua experiência na utilização do MAID.

☐ Divertida	☐ Fácil	☐ Precisa	☐ Útil	☐ Cansativa
☐ Intuitiva	☐ Imersiva	☐ Inspiradora	☐ Estimulante	☐ Aborrecida
☐ Irritante	☐ Irrelevante	☐ Complexa	☐ Inovadora	☐ Entusiasmante
☐ Intimidante	☐ Confusa	☐ Atractiva	☐ Motivante	☐ Desconfortável

5. Sugestões e Comentários

____

____

____

Muito obrigado pela tua colaboração.

Annex 3 – ACE 2012 Paper: A Gesture Interface Game for Energy Consumption Awareness

A Gesture Interface Game for Energy Consumption Awareness

Ricardo Salvador, Teresa Romão, Pedro Centieiro

CITI, DI-Faculdade de Ciências e Tecnologia/UNL
Quinta da Torre
2829-516 Caparica, Portugal
+351 21 2948536

rmssalvador@gmail.com, tir@fct.unl.pt,
pcentieiro@gmail.com

Abstract.
Decreasing the energy consumption is an important goal for environmental sustainability. This paper describes MAID (Motion-based Ambient Interactive Display), an interactive public ambient display system, driven to motivate behavior changes regarding domestic energy consumption, through a persuasive game interface based on gesture recognition technology. The developed prototype guides players through the different rooms of a house, where they have to find out what is wrong and practice the correct actions to save energy, using similar gestures to the ones they would use in the real world to achieve the same goals. The system provides feedback regarding the consequences of each action. The paper also describes and presents the results of a user evaluation study performed during an open day event, attended by 6000 high school students, at our University Campus.

Keywords: Persuasion, Behavior change, Public Ambient Displays, Kinect, gesture interfaces.

1 Introduction

Many environmental problems occur due to our erroneous and exaggerated use of Earth resources. Recent studies suggests that, within USA, 22% of domestic energy consumption could be eliminated, if people engage in more appropriate behaviors, such as replacing traditional light bulbs with low-energy or adjusting thermostats and turning off lights when leaving rooms [13]. Each one of us can contribute to reduce domestic energy consumptions and consequently CO_2 emissions. Small individual actions may seem insignificant, but together everyone's efforts can have a great impact on the environment. Persuasive technologies [5] can be used to make people aware of the consequences of their actions and teach them how to proceed, helping them to change their attitudes and behaviors. MAID is a game deployed on an interactive public ambient display (PAD) system, which explores the use of natural hand

adfa, p. 1, 2011.

gestures. The system delivers a persuasive interface, using gesture recognition and presence detection. The main objective of MAID is to instruct (or remind) users about simple procedures to save energy, showing them how easy it is to have a huge impact on the environment by taking simple actions in their daily lives. MAID provides users with an interactive scenario, which can be manipulated simply by hand gestures. Users need to browse the scenario, trying to find out what is wrong and applying the corresponding changes. The impact of those changes is explained by audio feedback and it can also be visualized by the corresponding saved amount of money and reduction in CO_2 emissions. MAID is intended to be deployed in public areas, persuading users that walk by, to learn and adopt pro-environmental behaviors in their everyday lives.

This paper follows the previous publication related to MAID [25], offering further and updated details on the inner workings of the prototype and presenting the results of user tests.

MAID has been developed in the scope of Project DEAP, which aims to introduce new paradigms for environmental awareness, helping to motivate citizens to become more environmentally responsible in their everyday lives and engaging them in environmental preservation activities.

This paper is structured as follows: section 2 describes the state of the art related to the main areas covered by the presented work; section 3 presents MAID prototype; section 4 deals with the planning, realization and results of the user tests, and finally, section 5 presents conclusions and directions for future work.

2 Related Work

Persuasive technology is the general class of technologies that purposefully applies psychological principles of persuasion to interactive media, aiming at changing users' attitudes and behavior [11]. The term captology was coined by BJ Fogg, and is derived from computers as persuasive technology [5]. It focuses on the design, research, and analysis of interactive computing products created to change people’s attitudes and behaviors.

Persuasion technologies can and have been applied within several contexts, such as commerce, health care [3], education [17] and environment [4, 7, 16, 21, 15]. Several approaches explore the use of games as persuasive tools to influence players to change certain behaviors. Regarding domestic energy consumption, some applications have been developed. Power Agent [1] is a mobile pervasive game for teenagers, designed to influence their everyday activities and electricity usage patterns in the household. Power Agent is played on standard Java-enabled mobile phones and uses real power consumption data collected via existing metering equipment used by the energy company. The underlying design idea is to let players (one for each house) compete in teams (cooperate with their families) and learn hands-on how to save energy in their homes. Power Explorer [8] builds on the previous research on Power Agent and uses a similar technological set-up, but focuses on real-time feedback. The game design teaches the players about the consumption of their devices and encourages them to adopt good habits. Some other applications focus on providing feedback based on real-time domestic energy consumptions measurements [10, 14, 18].

Although these approaches seem to produce positive results, they rely on the use of mobile devices and the deployment of specific energy consumption devices (connected to the home central electric power meter or appliances), which involves additional costs and set-up. We aim at using players' free time in public spaces, such as shopping centers or waiting rooms, to both motivate and educate them towards home energy saving behaviors. Our approach is based on the assumption that a change in behavior requires both motivation to act as well as knowledge of what to do. According to Fogg [6], behavior changes can be achieved through the conciliation of three main factors: Motivation, ability and trigger. Thus, our design exposes several common inappropriate behaviors that must be noticed and corrected by the players through a public display. Actions that can be taken to reduce energy consumption include turning lights off, unplugging standby equipment, as well as replacing traditional light bulbs with low-energy ones. The interaction with the public display is performed by hand gestures similar to the ones required to perform the appropriate action in real life, so players can learn and rehearsal these actions and may easily recall them later. Additionally, to make players aware of the consequences of their actions, the system provides feedback regarding the costs saved and the quantity of CO2 emissions avoided. To have a more impressive effect and emphasize how individual actions can have a significant contribution to the global achievements, feedback is based on what would be saved or avoided if everyone in the country would take a certain action.

Hand gestures were chosen to avoid the need for any additional device, as well as to allow players to perform actions as similar as possible to the real ones. According to [9] gestures are a viable method of interaction with public digital. Their work contributes to the understanding of how people perceive, respond to and interact with interactive gesture based public displays outside the controlled environment of a research lab. Their findings support both the observe-and-learn model [19] and the Honeypot effect [2] and we followed some of their recommendations that were appropriate for our case study. A public display allows for a broader dissemination of a message, since besides the user directly interacting with it, all members of the audience can receive the output from the public display.

We also considered the four-phase framework" proposed by Vogel and Balakrishnan [23] which covers a range of activities from distant implicit public interaction to personal interaction, motivating users to get close enough to interact directly with the public display.

Diverse technologies have been explored to recognize gestures for interacting with public displays. Vogel and Balakrishnan [24] explored the use of free-hand interaction by employing a sensor data glove. Shoemaker et al. [20] introduced the shadow reaching technique that uses a perspective projection of the user's shadow on the display surface, controlled directly by the user through body positioning, to allows him to adjust the effective range of interaction over the a large area. However, Kinect sensor of the Microsoft Xbox console [11] represent a non-invasive technique for acquiring hand gestures which allows the creation of a natural interface that do not require user manipulation of additional devices.

Implementing a successful gesture-based interface also requires the challenging task of designing a set of effective gesture commands that allows users to interact in a natural, familiar and effortless manner. This is particularly important in the context of sporadic interaction with public displays when users have limited time to learn and explore the application [22]. As we mentioned before, our gesture set was designed to resemble the corresponding real world actions, so the gestures would be easier to understand and recall.

3 System Description

MAID (Motion-based Ambient Interactive Display) is a game deployed on an interactive public ambient display (PAD) system, which explores the use of natural hand gestures. It is driven to motivate behavior changes regarding domestic energy consumption, through a persuasive interface based on gesture recognition technology. The main objective of MAID is to instruct (or remind) users about simple procedures to save energy, showing them how easy it is to have a huge impact on the environment by taking simple actions in their daily lives.

The developed prototype guides players through the different rooms of a house, where they have to find out what is wrong and perform the correct actions to save energy, using similar gestures to the ones they would use in the real world to achieve the same goals. The impact of those actions is explained by audio feedback and can also be visualized by the corresponding saved amount of money and reduction in CO2 emissions. MAID is intended to be deployed in public areas, capturing the attention and persuading passing-by users to learn and adopt pro-environmental behaviors in their everyday lives.

It would be expected from a Public Ambient Display System that the main component would be the display itself. In most common implementations of a PAD, in order to add interactive components to the interface, the interaction method used is integrated in the display itself (touch-screen), or attached to it, in the form of button panels or controllers. To the common user, the main component will still be the display, but the center piece of the MAID's architecture is a motion-detection 3D sensor device - “Kinect for Xbox 360” by Microsoft - that is used to obtain the ambient contextual data. MAID is implemented in C++, using OpenGL technology for graphics processing, following a highly modular and configurable development strategy and allowing gameplay scripting. It can also be connected to a Facebook application, which allows users to share their experiences and increase the exposure of the project.

3.1 Architecture

The MAID development was done on a Linux based operating system, since early on most of the software development involving the Kinect was only available for this platform. So most of the technological options for the development process were based on what the Linux platform had to offer. The application is built over the NITE middleware, using the FreeGlut implementation of the OpenGL framework for graphics processing, being one of the most widely supported 3D graphic platform available for Linux. The textures applied in the application are loaded from png files, allowing transparency support. For configuration file management, the development was supported by the libconfig Linux library, following a modular and configurable development strategy in order to allow gameplay scripting. Most of the interface components, like all the textures used for the graphic rendering, the size and type of the fonts applied, the sound tracks used, the interaction parameters, the textual content, among many other options of the interface are configurable. The gameplay itself can be modified, by adding or removing rooms from the scenario, as well as by adding or removing interactive elements from each room.

Another component of the MAID is a sound engine that serves as a complementary mean of communication, speaking out loud most of the textual information presented in the interface, via Festival text-to-speech engine. It also manages the playback of soundtracks using the FMod framework, supporting multichannel direct playback of multiple audio formats, like wav and mp3. Figure 1 shows MAID's architecture.

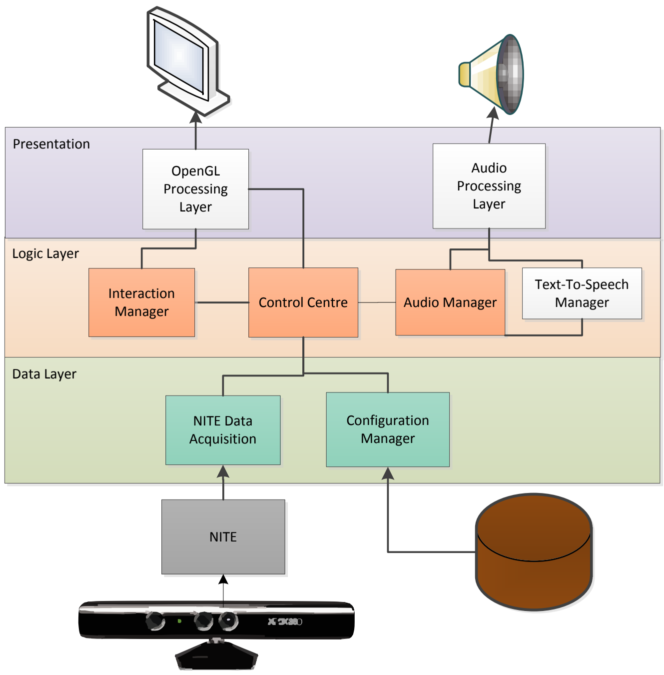


Fig. 1. MAID's architecture.

3.2 Gameplay

To play the game, users need to approach the public display and interact with the system by hand gestures. Several rooms of a house can be explored. Users should select one room at a time and take the appropriate actions to reduce energy consumption and CO2 emissions. The objective of the game is to save as much energy as possible, solving all the situations presented by the different game scenarios. Each situation is associated with a different electronic equipment that, when misused, can decrease the home energy efficiency.

Players receive visual and audio feedback while performing the different actions to save energy, as explained later in the paper, to make them aware of the consequences of their actions and how their small contribution can produce a significant impact in the global reduction of CO2 emissions.

3.3 Interface

To motivate users to get close enough to interact directly with the system's information and according to [23], a four stage interface was created:

- Stage 1: An idle state, in a context where no user was identified in the surroundings. The interface presents minimal content.
- Stage 2: A user was detected, and the interface tries to captivate his attention through a greeting message.
- Stage 3: A user is still standing in front of the display and the application invites him to further interact with the system. The user is invited to produce an activation gesture, with his hand, in order to calibrate the gesture recognition system.
- Stage 4: An interactive scenario is presented to the user. He can now explore it, through hand gestures.

To exit the game, the user just has to walk away from the PAD.

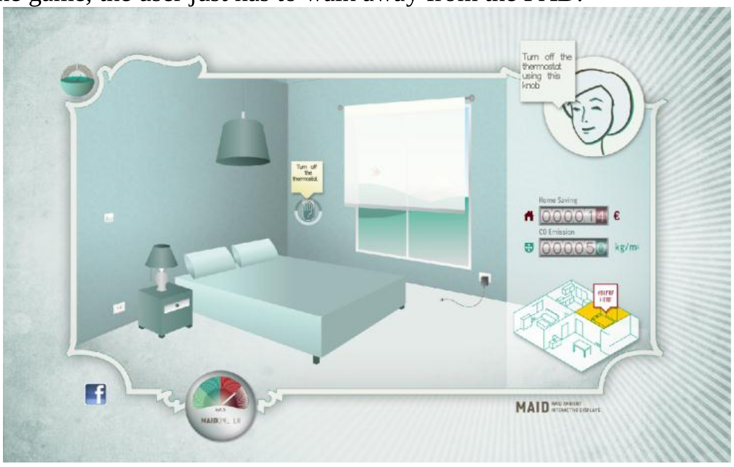


Fig. 2. The MAID's main interface.

The first three stages are articulated in order to build momentum and anticipate the final stage. In this final stage a typical video game interface format is presented, composed of a game terrain, a head-up display (HUD) and a pointer (Figure 2). As explained below, it also features a direct interaction mode, which allows users to perform actions, like changing a lamp, by using hand gestures.

Game Terrain.

The game terrain is a 2D representation of a home scenario. The user will start in a room of the house, and at any given time, he can select another room from the house plan in the lower right corner of the display. Each room has a set of interactive areas/objects, in a gameplay model similar to what could be found in a point-and-click adventure game. These areas correspond to situations that should be corrected, representing desirable user behavior changes. The players should scan the scenario and search for these situations, controlling the pointer (see 3.3.3) through hand-gestures. Some clues are provided by the system: pop-up balloons point out each situation and an assistant speaks out a description of the situation when the corresponding area in the game terrain is appointed by the user.

HUD

The HUD is composed by a configurable set of gages translating the current game status into a graphical representation that the player can easily assess (figure 2):

- A classic needle gage showing the current consumption of the home in Kw/h.
- A counter gage presenting the daily CO^2 emissions that would be avoided if the actions were taken in every home in the country, in t/m^2.
- A counter gage showing the home annual costs saved when taking the current actions, in euros.

These gages help to deliver the persuasive messages, urging the users to realize that together small individual behavior changes can have a significant impact in the environment and to change their behaviors in real life. In the HUD there is also an interactive house plan that allows the user to move to a different room in the house. The final component is the MAID assistant that delivers information to the user when the context demands so, in the form of a pop-up speech balloon that is also read through a text-to-speech system. The MAID assistant gives instructions that help the users to complete tasks and hints leading the users to accomplish the objectives of the game.

Pointer.

The pointer is directly controlled by the hand of the user during the exploration of the 2D scenarios. In each scenario, there is a set of hoverable areas. When the pointer hovers over one of these areas, a circular progress bar pops up, representing the hovering elapsed time. While hovering, a tooltip pops up with information related to the actions that could be taken to optimize that particular situation (figure 2). The MAID assistant, in the HUD, also delivers a speech regarding the hovered area. If the hovering lasts for the full load time, the user then enters the direct interaction mode.

Fig. 3. Direct Interaction Mode: Turning off a Thermostat.

Direct Interaction Mode.

When entering the direct interaction mode, the interface changes: the HUD disappears, and the center stage is cleared to house an enlarged version of the chosen situation/object (Figure 3). Here the user is asked to perform a specific hand gesture, which triggers a reaction on the object. For example, the user could enter the direct interaction mode by selecting a Thermostat in the game scenario (Figure 2), he would then be asked to perform a circular gesture to turn off the thermostat (Figure 3). The correct gesture is illustrated by the ever present MAID assistant (an animation depicts the correct gesture) and it is intended to match the gesture that should be done to perform the corresponding action in the real world. Upon the completion of the challenge, a congratulatory message is displayed along with complementary information, like fun facts, related to the action performed. Finally the interface returns to its previous state, allowing the user to explore another situation/object. At this point the counter gauges are updated to reveal how much would be saved in costs and CO2 emissions if this action was performed in every house in the country. To make these changes in the counter gauges noticeable, the sound of changing gauges is played to match the gauge animation resembling a classic split-flap display.

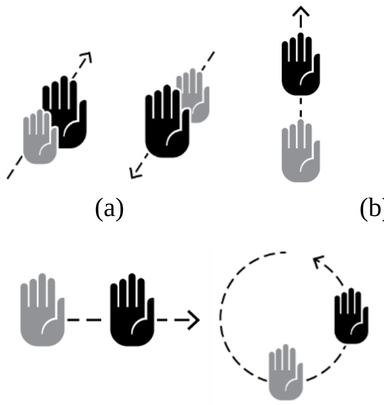
Gesture Commands.

There are four distinct gesture types that are implemented in MAID and can be used to create new interactive objects. The gesture recognition is based on single hand position detection. The length of the hand movements is configurable, but generally it should not go beyond an arm's reach, so the user doesn't have to move around to produce the target gestures. This is done to reduce the complexity of the gestures, to reduce the clutter free area needed for the detection and also to increase the reliability of the hand position recognition.

The following gesture types were implemented (figure 4):

- Pushing or pulling hand motion, used for objects' manipulation such as pulling a power adapter from a wall socket or tapping a light switch.
- Vertical sweeping hand motion, used for objects' manipulation, such as opening window shutters.
- Horizontal sweeping hand motion, used for objects' manipulation, such as opening sliding windows.
- Circular hand motion, used for objects' manipulation, such as unscrewing a light bulb and screwing in a new one or adjusting a thermostat.

The gesture commands were designed to resemble the corresponding real world actions, so they could provide a more natural interface and be easier to understand and to recall later.



(a) (b)

(c) (d)

Fig. 4. Gesture commands: (a) Pushing and pulling; (b) Vertical Sweeping; (c) Horizontal Sweeping; (d) Circular motion.

Facebook Interaction.

At any given time, while exploring the game terrain, the user can publish a picture with the current game status to the Facebook social network, sharing them with the world. The interface that guides the user through the process is based on the direct interaction mode, where a hoverable Facebook icon is integrated in the interface (figure 2), outside the frame that contains the game terrain. After hovering over the icon the interface moves to a mode similar to the direct interaction mode, where a mirror image of the Kinect RGB camera feed is presented, with a circular progress bar painted over. While the circle fills in the user has time to strike a pose and get ready for the shot. As soon as the circle is fulfilled (Figure 5), a picture is taken and the user is informed that the picture was published in the project DEAP's Facebook wall. In order to send the information to the social network, a Facebook Application was created, and together with the association of the application with the DEAP project Facebook page, access tokens were generated to allow us to publish the pictures directly from the MAID to the DEAP Facebook page photo gallery. This mechanism is implemented using a cURL command that sends the pictures and the corresponding descriptions to the photo gallery, using the generated access token for authentication. The picture is published in the MAID gallery of the DEAP project Facebook page, with a description that reflects the player's current progress in the game, for example (Figure 5):

"Action-shot taken by the MAID: At the time of this picture, the user had optimized the power consumption of the MAID home, in order to save up to 64Kw per month."

Fig. 5. Facebook picture capturing interface.

Fig. 6. MAID photo gallery picture.

In this case the user had optimized the MAID home in order to save 64Kw per month. After this data is published, the result isn't presented to the user, in order to persuade him to visit the project's Facebook page (figure 6), where he can "like" or even "tag" any picture, exploring the social network graph structure to spread the information and divulge the message behind the MAID.

4 Evaluation

A user study was conducted to evaluate MAID usability, gameplay and effects on users. Twenty six users aged 15-28 (average of 17.5) participated in the evaluation procedures (21 male and 5 female). All of them were familiarized with new technologies.

MAID was deployed on a large event for teenagers, an open day at our University Campus attended by 6000 high-school students. They were free to participate in the MAID evaluation as well as to experiment other activities in the event.

The MAID has three basic factors to take into account for the setup process: space, lighting and audience. Regarding space, the Kinect needs the player to be at least 1.8m away from the sensor and a 1.8m wide area around the player that is clutter free [11]. Concerning the audience, the display system used has to take into account the size of the potential audience, so for a large audience we use a projector, but for a small audience we can use a large wide-screen display. Regarding lighting, regular projectors need low lighting to deliver an acceptable definition and also the Kinect sensor should not be exposed to direct sunlight. For this evaluation study we had enough space and control over lighting conditions, which allowed us to use a projector and get a projection area large enough so that the audience does not need to crowd the user interacting with the MAID. The projector was overhead mounted to maximize the unobstructed area needed for the projection. The Kinect was mounted below the projection screen, and the users were positioned at a minimum distance of 2 meters from the sensor, with about 1.8m radius with no obstruction, allowing freedom of movements. Due to space constraints (other activities were occurring in the same room), audience was kept on the right side of the MAID deploying area.

The test sessions were conducted by two researchers, who played the roles of facilitator and observer. The first one had a more active role, giving an initial briefing and instructions to the participants and providing assistance for any problems that users might face. The second researcher focused on observing the way the tests unfolded, and how users reacted and interacted with the system.

Before starting to use the application, users were informed about the objectives of the test. After a short description of the application and an explanation of the goals to be achieved, users were encouraged to explore MAID. They could do it for as long as they liked, as in a real setting. During the tests the observer took notes of the users' behavior and verbalizations, providing a record of any issues explicitly mentioned by users.

At the end of each test, participants were asked to answer a questionnaire to assert their reactions, gather some personal information and collect their feedback and comments.

4.1 Questionnaire

The questionnaire captured users' personal data, experimental feedback and comments. The first part covered basic personal information like the participant's age, gender, familiarity with new technologies and frequency of use of Microsoft Kinect.

The second group of questions was related with general feedback and usability issues, as well as with the appropriateness of the gestures (figure 8 e 9). Users were asked to rate statements, using a five-point Likert-type scale, which ranged from strong disagreement (1) to strong agreement (5).

The third part of the questionnaire included questions about the users' energy consumption attitudes and how the experience motivated them to change their behavior. The fouth part of the questionnaire was based on the Microsoft "Product Reaction Cards", in order to capture the user's feelings towards the MAID. Users were asked to choose the words that best described their experience while using the application from a list of 20 words, a set that included about 60% of words considered positive and 40% considered negative. The participants could choose any number of words that they deemed adequate to describe their reaction. Finally, the questionnaire also included an open question, in order to gather comments and recommendations regarding future developments of additional features and to obtain a more general evaluation of the system.

4.2 Results and Discussion

Many participants had never used the MS Kinect before (54%), as shown in figure 7, and MAID had took on the rule of introducing them to gesture based interfaces.

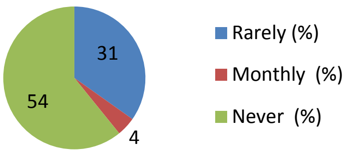


Fig. 7. Kinect usage.

Regarding the second group of questions, the results are very positive, as it can be observed in figure 8 and 9. The majority of participants liked to use MAID and found it easy to learn and to use. They considered the feedback and the information provided by the application useful and they were willing to use MAID in public spaces. Observations revealed that users had fun while using MAID and their friends, in the audience, were commenting and giving suggestions.

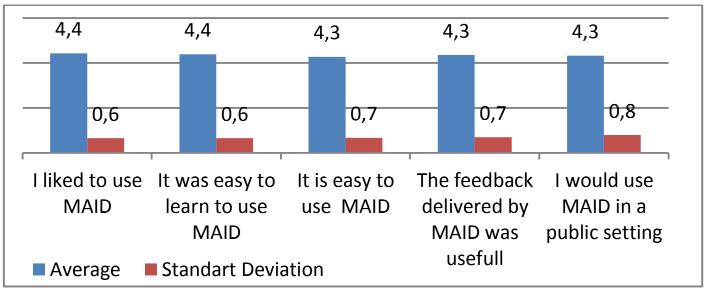


Fig. 8. Summary of results from general feedback and usability.

Results regarding the gesture commands were also very positive (figure 9). Participants considered the gestures natural and appropriate to mimic the real action, as well as easy to perform. Horizontal sweep was the gesture that marginally achieved the most positive results.

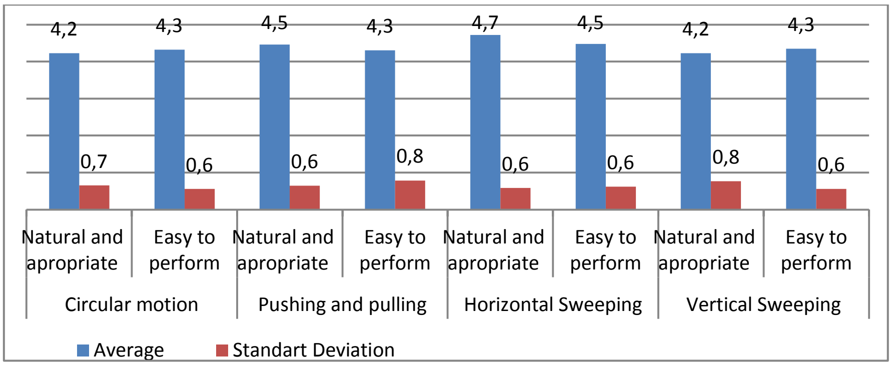


Fig. 9. Summary of results regarding gestures commands.

The questionnaire also included questions about the users' energy consumption attitudes and how the experience motivated them to change their behavior. All participants except two, considered that MAID provided them with new useful information and knowledge regarding domestic energy consumption and that their experience using MAID made them aware of the consequences of their appliance usage patterns. Half of the participants were already following some of the MAID suggestions in their everyday lives before the user tests. But most importantly, 50% of the participants reported that they were motivated by MAID to implement the given suggestions, and 46% answered "maybe". So, MAID seems to have a persuasive effect and may potentially influence people to change.

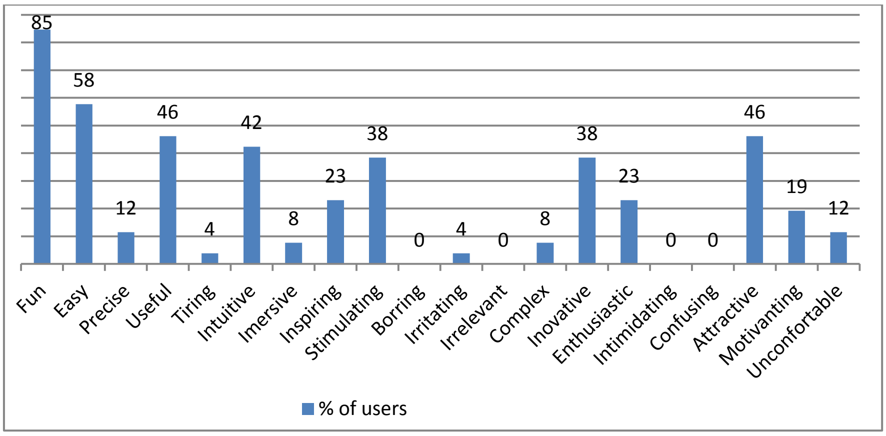


Fig. 10. – User experience feedback.

From the analysis of question based on the Microsoft "Product Reaction Cards" we concluded that all participants held positive feelings when classifying their experience using the MAID (figure 10), 85% of the users selected the word "Fun", which is an important quality of an application of this kind to be deployed in public spaces. Other positive top selected words were: Easy (58%), Attractive (46%) and Useful (46%). All positive words included in the set were selected by the users, and given the word choice, the users valued the overall look and usability aspects of the application. Even though most of the feedback was positive, some negative words were also selected by only one or two participants: Complex (8%), Tiring (4%) and Irritating (4%).

A few users reported that the hand gesture interface became tiring after a few minutes of interaction. Applications like MAID should not be used for a long time and users must be able to accomplish the game objectives in a few minutes.

The Facebook interface was also well received by the users, but only two participants took individual pictures. Most participants took group pictures with their friends. Fourteen photos were published, but most participants seem to forget to later visit the project Facebook page to check their pictures. A different mechanism must be explored to connect MAID with Facebook, preferentially without requiring users to login to their Facebook account when using MAID.

5 Conclusion and Future Work

This paper presents MAID, an interactive public ambient display system, driven to motivate behavior changes regarding domestic energy consumption, through a persuasive game interface based on gesture recognition technology.

The system was evaluated during a large event for teenagers, an open day at our University Campus attended by 6000 high-school students. Results were very positive. Most participants in the user tests liked to use MAID, found it fun and easy to use and considered that MAID provided them with new useful information regarding domestic energy consumption and that they were motivated to implement the suggestions provided by MAID in their everyday lives. Therefore, MAID seems to have the potential to influence people to change. However, we are aware that further tests need to be conducted in order to measure the persuasive effect of the game in the long term.

According to the evaluation results, the designed gesture commands seem to be appropriate for the purpose and easy to perform.

The Facebook interaction component of MAID would also benefit from some improvements, like giving the user the possibility to confirm or retake a shot, before it is sent to Facebook, or even cancel the whole process without taking any shot. This would increase the usability of this functionality. It also needs a mechanism that would allow the user to link the photograph taken directly to his account. This could be done with a QR code, were the MAID after taking the shot and sending it to Facebook, would generate a unique QR CODE that would link to the picture, so that the user could use a handheld device, such as his mobile phone to open the link and see the picture through his Facebook account.

6 Acknowledgments

This work is funded by Fundação para a Ciência e Tecnologia (FCT/MEC), Portugal, in the scope of project DEAP (PTDC/AAC-AMB/104834/2008) and by CITI/DI/FCT/UNL (PEst-OE/EEI/UI0527/2011). The authors thank Bárbara Teixeira for her contribution on the graphic design.

7 References

1. Bang, M., Gustafsson, A. and Katzeff, C.: Promoting new patterns in household energy consumption with pervasive learning games. In Proc. of Persuasive 2007 (Palo Alto, USA, April 26-27, 2007). LNCS, vol. 4744, pp. 55-63. Springer Verlag, Heidelberg (2007).
2. Brignull, H and Rogers, Y.: Enticing People to Interact with Large Public Displays in Public Spaces. In Proc. of INTERACT'03 (Zurich, Switzerland, September 1-5, 2003), pp. 17-24 (2003).
3. de Oliveira, R., Cherubini, M., and Oliver, N.: MoviPill: Improving Medication Compliance for Elders Using a Mobile Persuasive Social Game. In Proc. of UbiComp 2010 (Copenhagen, Demark, September 26-29, 2010), pp. 251–260. ACM, (2010).
4. Dillahunt, T. ,Becker, G., Mankoff, J. and Kraut, R.: Motivating environmentally sustainable behavior changes with a virtual polar bear. In Proc. of Pervasive 2008 Workshop on Pervasive Technology and Environment Sustainability (Sydney, Australia, May 19-22, 2008), pp. 58-62 (2008).
5. Fogg. B.J.: Persuasive Technology: Using Computers to Change What We Think and Do. Morgan Kaufmann Publishers, San Francisco (2003).
6. Fogg, B.: A behavior model for persuasive design. In Proc. of the 4th International Conference on Persuasive Technology (Claremont, USA, April 26-29, 2009), 40:1-40:7 (2009).
7. Froehlich, J., Dillahunt, T., Klasnja, P., Mankoff, J., Consolvo, S., Harrison, B. and Landay, J.: UbiGreen: Investigating a Mobile Tool for Tracking and Supporting Green Transportation Habits. In Proc. of CHI '09 (Boston, USA, April 4-9, 2009), pp. 1043-1052. ACM Press, New York (2009).
8. Gustafsson, A., Bang, M., and Svahn, M.: Power Explorer: a casual game style for encouraging long term behavior change among teenagers. In Proc. of ACE 2009 (Athens, Greece, October 29-31, 2009), pp. 182–189. ACM, New York (2009).
9. Hardy, J., Rukzio, E. and Davies, N.: Real world responses to interactive gesture based public displays. In Proc. of the 10th International Conference on Mobile and Ubiquitous Multimedia (MUM 2011), Beijing, China, December 7–9, 2011, pp. 33-39 (2011).
10. Jahn, M., Jentsch, M., Prause, C. R., Pramudianto, F., Al-Akkad, A., and Reiners, R.: The Energy Aware Smart Home. In Proc. of FutureTech 2010, IEEE Press, NY (2010).
11. Kinect for XBox 360, http://www.xbox.com/kinect (2011)
12. Kort, Y., Ijsselsteijn, W., Millden C. and Eggen, B.: In Proc. of Persuasive 2007 (Palo Alto, USA, April 26-27, 2007). LNCS, vol. 4744, Preface. Springer Verlag, Heidelberg (2007).
13. Laitner, J.A., Ehrhardt-Martinez, K., McKinney, V.: Examining the Scale of the Behaviour Energy Efficiency Continuum. European Council for an Energy Efficient Economy Conference, June 1-6, Cote d'Azur, France, pp. 217-223 (2009).

14. Mattern, F., Staake, T., and Weiss, M.: ICT for green: how computers can help us to conserve energy. In Proc. of e-Energy '10 (2010).
15. Miller, T. M., Rich, P. and Davis, J.: ADAPT: audience design of ambient persuasive technology. In Proc. of CHI EA '09, pp. 4165-4170. ACM, New York (2009).
16. Reitberger, W., Ploderer, B., Obermair, C., and Tscheligi, M.: The PerCues Framework and its Application for Sustainable Mobility. In Proc. of Persuasive '07 2nd (Palo Alto, USA, April 26-27, 2007), pp. 92-95. Springer-Verlag, Heidelberg (2007).
17. Revelle, G., Reardon, E., Green, M., Betancourt, J. and Kotler J.: The Use of Mobile Phones to Support Children's Literacy Learning. In Proc. of Persuasive '07 (Palo Alto, USA, April 26-27, 2007), pp. 252-258. Springer-Verlag, Heidelberg (2007).
18. Reeves, B., Cummings, J. J., and Anderson, D.: Leveraging the engagement of games to change energy behavior. In Gamification Workshop at CHI 2011 (2001).
19. Rubegni, E, Memarovic, N, and Langheinrich, M.: Talking to Strangers: Using Large Public Displays to Facilitate Social Interaction. In Proc. of the 14th International Conference on Human-Computer Interaction (HCII 2011), Orlando, USA, July 9-14, pp. 195-204 (2011).
20. Shoemaker, G., Tang, A., and Booth, K. S.: Shadow reaching: a new perspective on interaction for large displays. In Proc. of UIST '07, pp. 53-56 ACM, New York (2007).
21. Takayama, C., Lehdonvirta, V., Shiraishi, M., Washio, Y., Kimura, H. and Nakajima, T.: ECOISLAND: A System for Persuading Users to Reduce CO2 Emissions. In Proc. of the 2009 Software Technologies for Future Dependable Distributed Systems (STFSSD '09), pp. 59-63. IEEE Press, Washington (2009).
22. Vatavu,R. D.: Nomadic Gestures: A Technique for Reusing Gesture Commands for Frequent Ambient Interactions. Journal of Ambient Intelligence and Smart Environments 4 (2), 79-93 (2012).
23. Vogel, D. and Balakrishnan, R.: Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In Proc. of the 17th Annual ACM Symposium on User interface software and technology (UIST '04), pp. 137-146. ACM, New York (2004).
24. Vogel, D. and Balakrishnan, R.: Distant freehand pointing and clicking on very large, high resolution displays. In: Proc. of the 18th Annual ACM Symposium on User interface Software and Technology (UIST'05), pp. 33-42. ACM, New York (2005).
25. Salvador, R and Romão, T.: Let's move and save some energy. In *Proc. of the 8th International Conference on Advances in Computer Entertainment Technology* (ACE '11), Article 86. ACM, New York (2011).