9-12-2012 12:00 AM

Interactive Augmented Reality As A Support Tool For Parkinson's Disease Rehabilitation Programs

Andrés Ayala García
The University of Western Ontario

Supervisor
Dr. Michael James Katchabaw
The University of Western Ontario

Graduate Program in Computer Science

A thesis submitted in partial fulfillment of the requirements for the degree in Master of Science

© Andrés Ayala García 2012

INTERACTIVE AUGMENTED REALITY AS A SUPPORT TOOL FOR PARKINSON'S DISEASE REHABILITATION PROGRAMS

(Spine Title: Augmented Reality for Parkinson's Disease Rehabilitation Programs)

(Thesis Format: Monograph)

by

Andrés Ayala García

Graduate Program in Computer Science

Submitted in partial fulfillment
of the requirements for the degree of
Master of Science

School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario
September, 2012

THE UNIVERSITY OF WESTERN ONTARIO
SCHOOL OF GRADUATE AND POSTDOCTORAL STUDIES

CERTIFICATE OF EXAMINATION

Supervisor   Examiners

_____   _____

Michael James Katchabaw   Hanan Lutfiyya

                                              _____

Supervisory Committee   Mahmoud El-Sakka

_____   _____

                                              David Walton

The thesis by
Andrés <u>Ayala García</u>
<u>entitled</u>

INTERACTIVE AUGMENTED REALITY AS A SUPPORT TOOL FOR
PARKINSON'S DISEASE REHABILITATION PROGRAMS

is accepted in partial fulfillment of the
requirements for the degree of
Master of Science

_____   _____
Date   Chair of Thesis Examining Board

Abstract

In this thesis, an augmented reality system is proposed as an alternative to create multiple interactive virtual environments that might later be used in Parkinson's Disease rehabilitation programs. The main objective of this thesis is to develop a Wearable Tangible Augmented Reality Environment focused on providing the sense of presence required to effectively immerse patients so that they are able to perform different tasks in context–specific scenarios. By using our system, patients are able to freely navigate different virtual environments. Moreover, by segmenting and then overlaying users' hands and objects of interest above the 3D environment, patients have the ability to naturally interact with both real–life items as well as with virtually augmented objects using nothing but their bare hands.

As part of this thesis, Parkinson's Disease patients participated in a three–week dual–task assessment program in which several tasks were performed following a strict protocol. In order to assess patients' performance, the tasks were carried out both in the real world and using the system. The findings of this thesis will help evaluate the viability of using augmented reality as an auxiliary tool for Parkinson's Disease rehabilitation programs.

Keywords: Augmented Reality, Parkinson's Disease, Natural Interaction

Acknowledgments

I would like to thank Dr. Michael James Katchabaw for his support and guidance throughout the time I spent at the University of Western Ontario. Professor, thank you for believing in me and for giving me the opportunity to collaborate with you in this very interesting project.

I would like to thank Dr. Mandar Jog and Dr. Angela Roberts South for the guidance, teachings and support that you kindly provided during the development of this project. It has been an honor and a privilege to work with both of you.

I would like to acknowledge the financial support of the National Council of Science and Technology (Consejo Nacional de Ciencia y Tecnología) and the Electric Research Institute in Mexico (Instituto de Investigaciones Eléctricas) in Mexico.

I would like to thank my parents, because they have always been the source of my inspiration. Thank you guys for being so loving and the most amazing parents.

I would like to acknowledge all the wonderful people in Mexico and Canada that helped to make this adventure possible. Thank you to all my new and old friends for your support, kindness and the good moments we have shared together.

But above all, to my beautiful Wife, Ilse Leal Aulenbacher, thank you for being my mentor and my best friend. You are always the best companion in every journey we have embarked on together. Thank you for your loving support and for always brightening my days.

Table of Contents

List of Figures

Chapter 1

Introduction

1.1 Preliminary

Parkinson's disease (PD) is a progressive degenerative disorder of the central nervous system characterised by a large number of motor and non–motor features that can impact on function to a variable degree. There are four main motor features of PD: tremor at rest, rigidity, akinesia (loss of control of voluntary muscle movements) and postural instability [30]. Gait is one of the most affected motor characteristics. Gait abnormalities can cause loss of balance and a tendency to fall, which often causes serious injuries [2]. In addition, the non–motor symptoms associated with PD include autonomic dysfunction, cognitive/neurobehavioral disorders, as well as sensory and sleep abnormalities [30]. As the percentage of the elderly in the population grows, the prevalence of PD in North America is expected to double in the course of the next 20 years. There is an important economic burden caused by the disease [14]. In the United States alone, the annual economic impact of PD is estimated at $10.8 billion, 58% of which is related to direct medical costs [50][27]. Given the economic impact, the decrease in quality of life caused by PD, along with the predicted rise in

prevalence, there is a substantial need for new and novel methods of treatment and rehabilitation for PD.

Unfortunately, there is currently no cure for PD, but there is medication and various forms of therapy and rehabilitation designed to help manage symptoms and improve patient's quality of life. However, several issues with current approaches to rehabilitation of patients with PD have been reported [39], being the lack of task and context–specific rehabilitation programs the main issue. Benefits from rehabilitation have been often linked to context, and the in–clinic context is typically contrived or artificial and does not adequately capture real life scenarios, situations or challenges that patients face in a daily basis. Limitations of the in–clinic environment restrict the types of activities that can be made as part of rehabilitation programs [39]. In particular, scenarios that are potentially hazardous or dangerous, yet are part of daily life, cannot be supported in current rehabilitation programs.

Recently the interest in Virtual Environments (VEs) has grown in the PD research community due to the potential that comes through the use of VEs. Different scenarios can be simulated, providing whatever "context" is needed, while bypassing inherent limitations of the current clinic environment and ensuring safety regardless of the scenarios presented. Many different VEs can be created through virtual or augmented reality technologies.

In this thesis, we created three different virtual environments using augmented reality. These environments allow us to assess patients with PD while they perform dual–task activities. How well the patients perform in those activities will help us evaluate the feasibility and limitations of using augmented reality as a support tool in PD rehabilitation programs.

1.2   Augmented reality

Augmented Reality (AR) is the visual combination of real–time video streaming and computer generated 2D and 3D imagery. Opposed to the classic Virtual Reality (VR) paradigm in which users are immersed in an entirely simulated world, augmented reality allows users to stay connected with the real wold while creating the illusion of being in a different physical location. Furthermore, AR provides users with the ability to see and interact with objects that are not present in their surroundings [21]. According to Azuma et al. [5], augmented reality applications should meet the following three requirements: AR should be the mixture of video sequence and computer generated imagery, AR applications have to run in real time and virtual objects have to be properly aligned (registered) with real world structures.

In AR, computer generated graphics are overlaid into the user's field of view. For example, graphics can be used to a) add supplementary information or instructions about the environment, b) insert virtual objects, c) enhance real objects, or d) provide step–by–step visual aids that are needed for the execution of a task. In its more basic form, augmented reality overlays simple head up displays, images or text into the user's field of view. More complex AR applications display sophisticated 3D models rendered in such a way that lighting conditions, shadows casting and the simulation of occlusions appear indistinguishable from the surrounding natural scene. Figure 1.1 shows an example of a common AR system in which the video image is acquired, registered and augmented. In order to register the virtual cereal box in the image, the AR system derives tracking information from the video input. After rendering the registered 3D transformation, the real object can take any other appearance or even be transformed into a completely different object. Every pattern printed on the cereal box could represent a different figure or used to provide interactive buttons. This

Figure 1.1 A simple augmented reality example

a) Original video feed   b) Augmented scene

type of visualization is a powerful tool for exploring real world structures along with additional contextual information. For example, by augmenting textual descriptions or instructions, AR displays are able to provide semantics to real world objects that, by themselves, do not provide any information about how they should be used. Since virtual information does not necessarily have to follow real–world physical rules, AR is able to present a variety of non–natural effects in a real world environment. This ranges from the inclusion of simple textual information floating on the space, to the incorporation of complex fictional characters living in the real world environment [61]. Augmented reality is useful to create special effects such as the illusion of uncovering hidden objects or generating the impression of seeing through formerly occluded objects.

It is important to mention that the concept of augmented reality is not restricted to the sense of sight. AR can potentially apply to all senses, augmenting smell, touch and hearing as well. However, the work on this thesis is entirely oriented to visual augmented reality using a Head Mounted Display (HMD).

Figure 1.2 Diagram of a simple augmented reality system
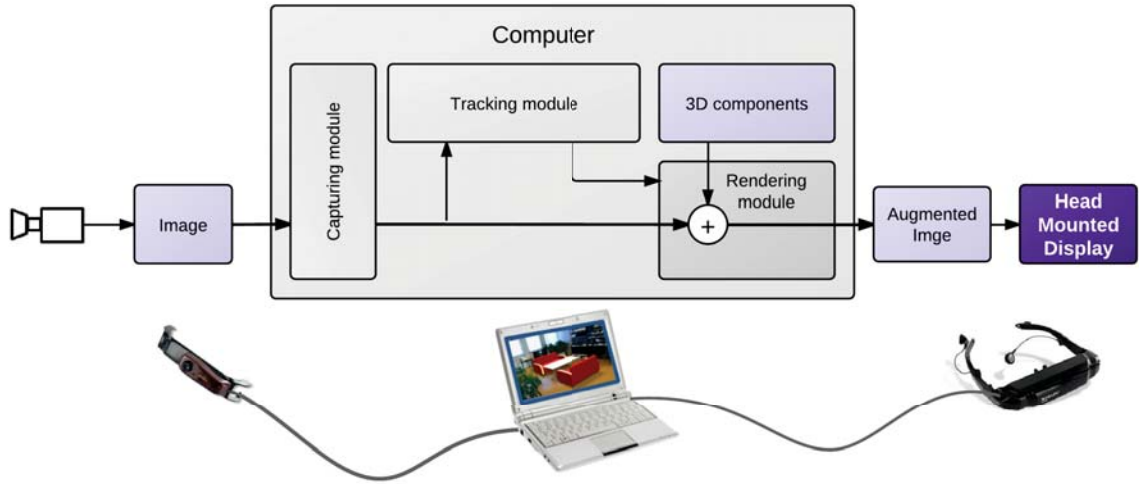
1.2.1 Registration and tracking

In order to appropriately integrate real and virtual information, both the real image and the 3D augmentation have to be carefully combined rather than simply attached together. If computer graphics are generated separately without correctly registering the visible real environment, a favorable visual composition between both types of data may not be accomplished. Providing robust and accurate *registration* is the main technical difficulty that AR systems have to overcome. In AR systems like ours, where head mounted displays are used, registration is equivalent to computing the pose (rotation and translation) of the user's viewpoint.

In AR, image registration uses video tracking algorithms that usually consist of two stages: *tracking* and *reconstructing*. In the first stage, fiducial markers or image features are detected. The tracking step usually employs feature detection, edge detection, or other image processing methods. The reconstructing stage uses the data

obtained from the first stage to reconstruct a real world coordinate system based on a camera model and object transformations [70]. Figure 1.2 shows a diagram that illustrates a simple AR system and its components.

1.2.2   Occlusion problem and depth perception

One of the inherent drawbacks of overlaying virtual environments to video, is that objects of interest are frequently occluded by 3D augmented objects, thus creating an unrealistic effect where foreground items that should appear in front of the augmented information are occluded (see Figure 1.3). Realistic image composition requires the correct combination between virtual and real objects, in which background/distant augmented objects must be correctly occluded by foreground real objects. Solving the occlusion problem in augmented reality is challenging when there is not enough information about the real world that is being augmented.

Figure 1.3 Unrealistic effects are created in cases in which augmented objects occlude real objects. In this picture the drawer should be rendered behind the chair.

If we do not take into consideration the information covered by the overlaying virtual objects, the resulting visualization may cause problems in depth perception. The human cognitive system interprets a set of monocular and binocular cues in order to interpret depth and spatial organization of the 3D objects in the environment [23]. Therefore, when carelessly adding virtual information to real–world video, the AR system may eliminate or alter some of those depth cues.

Monocular depth cues can be further divided into *pictorial cues* and *dynamic depth cues*. According to Furth [21], pictorial depth cues are those that can be found in a single image and include the clues listed below:

- **Occlusion**. If the 2D projections of two objects in the environment overlap, objects which are closer to the observer occlude objects which are further away.

- **Relative size**. More distant objects appear to be smaller than closer objects.

- **Relative height**. Objects with bases higher in the image appear to be further.

- **Detail**. Objects which are closer offer more detail.

- **Shadows**. Depending on the position of the light source, shadows can be cast from one object onto another.

Dynamic depth cues are caused by the fact that objects further away seem to move slower than objects that are closer to the observer when the viewpoint is moving. *Motion Parallax* is an example of dynamic depth cues. For the development of our system, we used a single camera mounted on a head mounted display. For that reason, we were mainly interested in simulating monocular pictorial depth cues.

1.3 Skin segmentation using color pixel classification

As mentioned in Section 1.2.2, in the classic augmented reality approach, what a user sees is a combination of two layers: video as background and 3D as foreground. One of the main challenges of augmented reality is the occlusion problem. In simple terms, occlusion is the process of determining which objects should be visible in relation to other objects. Occlusion provides a very important visual cue to the human perceptual system when rendering data in three dimensions [67].

For example, when we interact with the real world, it is clear that if we place our hand in front of some other object, for example a table, some part of it will be hidden by our hand. In augmented reality systems, occlusion is not always resolved successfully, leading to an unnatural and confusing experience for the user. Skin detection can help tackle this problem by identifying the set of pixels that correspond to skin in an image so that hands can be placed in a separate layer. Thus, instead of having two layers (video and 3D models), we are proposing the implementation of a third layer that would correspond to hands and other objects of interest. With the third layer, the occlusion problem can be corrected by placing skin pixels in front of both the 3D and video layers (Figure 1.4 shows a representation of the multilayer approach we are proposing to solve the occlusion problem). Machine learning algorithms can be of great aid for computer vision applications such as the implementation of a skin classifier. For this thesis, we implemented a two–class skin classifier using an Artificial Neural Network. A skin classifier is a two–class classifier that defines a boundary of the skin color class in the feature space [20]. In this context, the feature space corresponds to the way we choose to represent colors.

Figure 1.4 To avoid the occlusion problem (4), we overlay the 3D information (2) over the original image layer (1), and on top of the first two layers we add a third layer composed by the user's hands and other objects of interest (3). The result is a properly composed image (5).
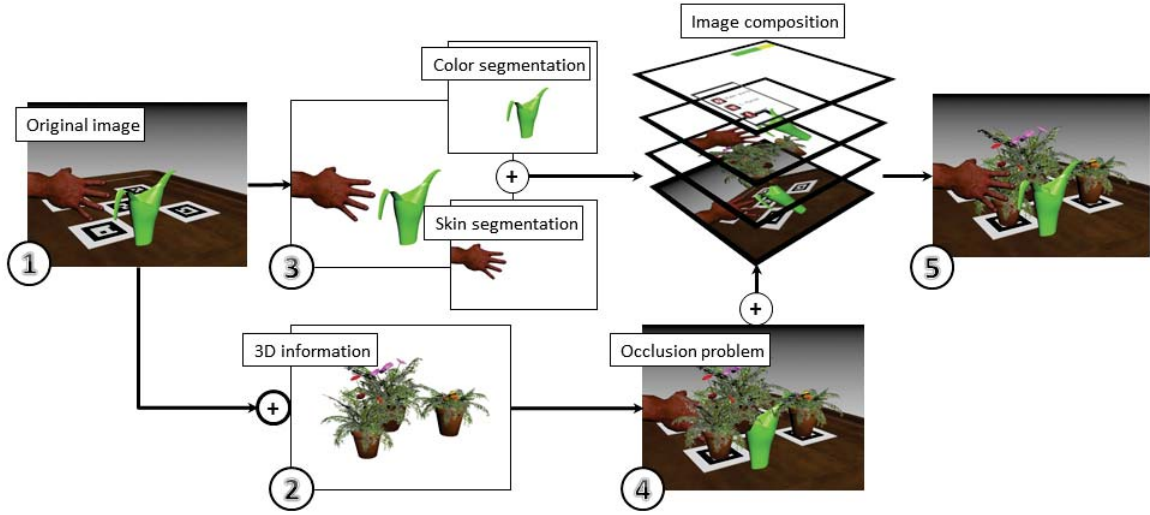
1.4 Presence

In virtual environments, *presence* can be defined as a state of consciousness, the psychological state of "being there" [56][24]. Witmer and Singer [66], defined presence as the subjective experience of being in one place or environment, even when people are physically situated in another.

A frequent example of presence has been the sensation experienced by teleoperators, which have reported the awareness of being at a remote work site rather than at the local control station [8]. If the concept of presence is applied to a virtual environment, presence refers to experiencing the tridimensional computer–generated environment

rather than the actual physical location. Involvement and immersion are two concepts of interest related to presence [66].

Depending on the extent to which people perceive different stimuli to be significant or meaningful, they can put more attention and become more involved in the virtual experience. Durlach et al. [18] remarked that attention and awareness that people focus to the set of stimuli or events generated by virtual worlds, can lead to an increased sense of presence. According to Witmer and Singer, immersion depends on the continuous flow of stimuli and experiences which are provided by a virtual environment, that make people feel included in, and capable to interact with the environment. Important factors that affect immersion are: how isolated users are from the physical environment (e.g. visually), the extent of perception of self–inclusion in the virtual environment, whether the system provides natural modes of interaction, and finally, the perception of self–movement.

One of the objectives of this thesis is to evaluate whether the proposed AR system provides the sense of presence required to virtually transport and immerse users inside the synthetic environment. Based on the work of Witmer and Singer [66], we asked patients with PD to answer a subjective presence questionnaire. The questionnaire was used to evaluate relationships among reported presence and other research variables. The results of such evaluations are described in Chapter 5.

1.5   Thesis contribution

The main goal of this thesis is to design, develop and evaluate a wearable augmented reality system, designed to assess patients with PD in dual tasking activities (per-

forming simultaneous motor and cognitive tasks). This system will allow patients with PD to interact with both augmented and real objects, using nothing but their bare hands. This approach is novel, because the system provides mechanisms to allow free and natural navigation inside a virtual environment. Through a head mounted display, patients with PD are immersed in 3D virtual environments. In this way, multiple context and task–specific scenarios can be represented. For instance, patients could be immersed inside a virtual environment representing a grocery store, in which they can perform tasks that commonly are difficult to patients with PD. Examples of such tasks are: bending over to pick something from a bottom shelf or walk through reduced aisles while avoiding virtual obstacles. This would allow physicians to observe the behaviour and the performance of patients as if they were present while patients do grocery shopping.

Previous VEs approaches in PD research have involved the use of virtual reality and expensive, faulty tracking devices. In our approach, a small video camera will be placed in front of a head mounted display. Video streaming will be used to visually track several fiducial markers that will be placed on strategic locations inside a room. Based on heuristics, the system will estimate the camera position and orientation relative to the room.

In order to solve the occlusion problem, we are proposing a three–layer rendering composition process, where the first layer is used to display the video streaming. The second layer contains all the 3D graphics and the third layer is used to superimpose objects of interest such as tools and users hands over the virtual information. Figure 1.4 shows the image composition process we are proposing.

Our system can be divided in four fundamental modules: `CoreSystem`, `VideoSource`, `ARDriver` and `ScenarioManager`. In the following sections, we will briefly discuss

each module.

1.5.1   CoreSystem: manager and container

This module provides methods to manage I/O events, callback functions and message passing between modules. `CoreSystem` creates all the instances of other main modules, as well as other auxiliary modules such as an XML parser and a 3D OBJ loader manager. CoreSystem also administrates the simulation main loop.

1.5.2   VideoSource: video capture and image processing

`VideoSource` uses OpenCV's highly efficient data structures and methods to process and expose video input to the other modules. OpenCV (Open Source Computer Vision Library) is an open–source BSD–licensed library that includes computer vision and image processing algorithms [1].The original video grabbed by VideoSource is used to create the first bottom layer which corresponds to the real world background scene. Additionally, `VideoSource` segments and extracts objects of interest from the video stream. The result of such segmentation is a binary mask that in a further process is used to generate the alpha channel of the upper third layer. The images generated by this module can be exposed to other modules through `CoreSystem`.

1.5.3   ARDriver: image registration and tracking

`ARDriver` contains methods for image registration, tracking and pose estimation. At every frame during the simulation, `ARDriver` recognizes the different markers and

computes their position based on hierarchical multi–marker configuration. All of the transformation matrices are exposed to other modules via `CoreSystem`.

1.5.4   ScenarioManager: rendering

Using OpenGL, `ScenarioManager` applies the matrix transformations computed in `ARDriver` to correctly project all the 3D models over the background image. After the 3D information is added, the third layer generated by `VideoSource` is alpha blended over the previous layers. `ScenarioManager` also provides methods to render the Graphical User Interface (GUI).

1.6   System evaluation

One of the objectives of this thesis is to evaluate the AR system performance. The system will be used in a series of trials. Those trials will follow a strict protocol approved by the University of Western Ontario Human Subjects Research and Ethics Board. The protocol and method for these trials will be described in Chapter 4. Following the protocol instructions, every patient is asked to perform several tasks, both in the real–world and using the AR system. The patients repeat the same set of activities in three appointments during three weeks. At the end of each appointment, they are asked to answer a questionnaire based on Witmer and Singer work[66]. The results of those questionnaires are used to evaluate if the system provides the sense of presence required for a more intuitive and immersive experience.

1.7   Thesis outline

Chapter 1 contains an introduction to augmented reality and Parkinson's Disease by explaining some basic concepts and the purpose of this thesis. In Chapter 2, we give an overview of related work. In Chapter 3, we describe the design and architecture of the proposed augmented reality system. The protocol and method of the study are defined in Chapter 4. In Chapter 5, we present the results of the trials we conducted to evaluate the performance of people using the system, as well as the results we obtained from the presence questionnaire. In Chapter 6, we present the limitations we found in the proposed system, the future work and conclusions.

Chapter 2

Related Work

Augmented reality is a multidisciplinary technology that, by its characteristics, shares concepts with many other disciplines including: virtual and mixed reality, human–computer interaction, video games, e–learning, computer graphics, computer vision, machine learning, cognitive psychology and perceptual psychology. In this chapter we present previous work on concepts and technologies that we consider to be meaningful and closely related to what we are proposing in this thesis.

Being patients with Parkinson's Disease the primary users of our augmented reality system, we believe that there is an inherent need to provide intuitive forms of interaction. In this chapter, we present different forms of interaction that other research groups have proposed in an attempt to allow users to select, manipulate and control virtual objects.

As we mentioned in Section 1.1, it is valuable to create task and context–specific rehabilitation programs for people with Parkinson's Disease. We are proposing the

creation of virtual scenarios to study if the performance of people doing dual–task activities is the same when immersed in virtual environments and in the real world. If performance is similar in both environments, it would be a good indication that patients with PD adapt well to AR and that they could benefit from context–specific rehabilitation programs performed in virtual scenarios. In order to assess the motor–cognitive interaction in Parkinson's gait, natural and unrestricted navigation is preferred. In this chapter, we present different metaphors and artifacts found in the literature that assess the navigation problem.

2.1 Registration and tracking

This section summarizes different tracking strategies used in augmented reality. Non–visual tracking technologies have been used in virtual environments. Active technologies that use magnetic fields or ultrasound are available. Some popular examples of magnetic trackers are the products produced by companies like Polhemus©[1]. InterSense©[2], produces inertial–ultrasonic hybrid tracking systems such as the IS–900[TM] system shown in Figure 2.2 (b). Even though commercial products are robust and provide low latency, they are not widely used in augmented reality due to their elevated cost. Moreover, they are still prone to errors caused by external factors such as interference. The low cost of video cameras and the increasing processing capacity of computers and handheld devices have inspired a significant increment of research into the use of video cameras as visual tracking sensors. The literature review in this section is focused on vision based tracking methods that have been used in augmented reality applications.

---

[1]http://www.polhemus.com/
[2]http://www.intersense.com/

In augmented reality, image registration uses different computer vision methods. Fiducial markers or interest points are detected from camera images. Tracking uses feature detection, edge detection, or other image processing algorithms to analyze live video from a camera. Tracking techniques can be divided in two classes: *feature–based* and *model–based*[71]. Feature–based algorithms consist of finding the relationship between 2D image features and their 3D world coordinates [40]. Model–based methods (see Figure 2.1) use real–world object heuristics. For example, a virtual model of tracked objects' features can be used. This virtual model is made available in memory as a Computer Aided Design (CAD) model of reference. Another example of a model–based method would be the use of 2D templates based on distinguishable features of an object. Once the relationship between the 2D image and 3D world

Figure 2.1 Model based tracking system loop from [37]. 1: Image acquisition, 2: Model rendering 3: Image measurement 4: Pose update
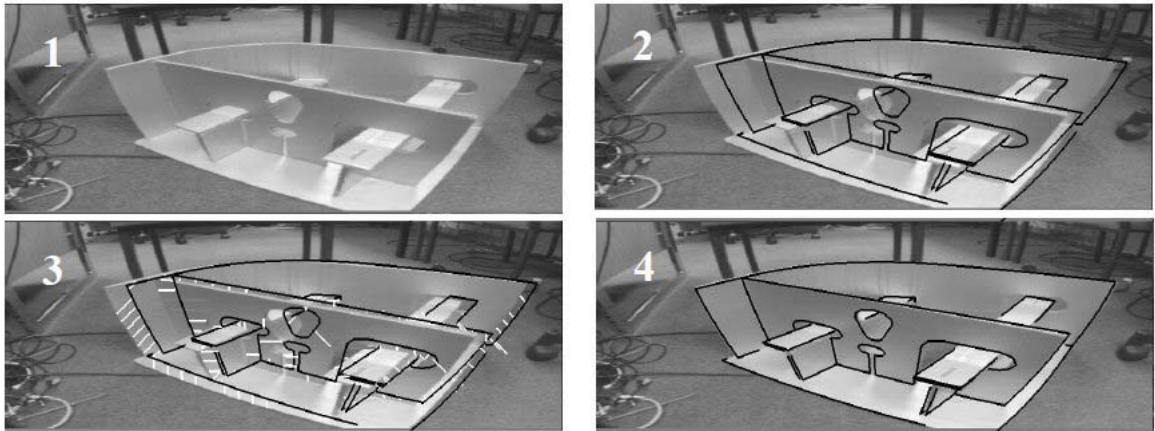
frame coordinates are found, the camera pose can be obtained by projecting the 3D coordinates of features into the observed 2D image. The reconstructing stage uses the data obtained from the first stage to reconstruct a real–world coordinate system.

Some methods assume the existence of fiducial markers in the surroundings. Other methods, like the one proposed by Huang et al., uses pre–calculated 3D structures for what they call the AR–View [26]. There are two important characteristics of the AR–view approach: the first one specifies that the camera has to remain stationary and the second one dictates that the position must be known beforehand. In their approach, when the scene is not known, they first use fiducial markers and Simultaneous Localization And Mapping (SLAM) to compute the relative position of the device with respect to the scene. AR–View has been used to virtually reconstruct the ruins of an important Chinese royal building called Yuanmingyuan. In cases where

Figure 2.2 Non–vision based commercial hybrid tracking systems

a) Polhemus PATRIOT wireless   b) IS-900 utrasonic tracking system

the AR device is static, an approach like the one adopted by Huang et al., can be used. On the other hand, if the AR device is mobile, tracking becomes much more difficult. Movable systems have to be able to model and deduce both camera motion and the structure of scene. For example in [49], a movable automotive pedestrian detector designed as a support system to prevent accidents was proposed by Nilsson et al. In that automotive system, they used augmented reality to add virtual agents to the video sequences. The augmented data allowed them to test their system in multiple modifiable scenarios that could have not been possible in real life.

There are some open–source AR libraries available for use, the most popular of them is ARToolKit. ARToolKit is a library that was developed based on the research of Hirokazu Kato from the Nara Institute of Science and Technology [34]. ARToolkit is a vision–based tracking library that uses real–time video to calculate the camera position and orientation relative to fiducial markers. Once the real camera position is known, the information can be used to correctly overlay 3D computer graphics over the markers. Many other libraries such as OSGART[3], NyARToolkit[4] and FLAR-Toolkit[5] were developed on top of ARToolkit. Additionally, two libraries (ARToolkitPlus and StudiersTube) emerged as successors of the original library. Both libraries added many features like the implementation of a class–based API; however, they are no longer being developed. Researchers have improved the capabilities of those libraries. For example, Owen et al., replaced the binary image inside the square rectangle with discrete cosine transform basis functions, improving the systems resilience to noise and occlusion [51].

2.2 Natural selection and manipulation

Many AR prototypes that support interaction are often based on classic desktop metaphors (for example, a mouse is needed to use on–screen menus, others require users to type on keyboards). Others make use of video game devices and controls such as joysticks, the Wii Mote®, PlayStation Move®, etc. Techniques popularized by handheld devices such as gesture recognition are also common in AR. The two main trends in AR interaction research are a) using heterogeneous devices to exploit

---

[3]http://www.osgart.org/index.php
[4]http://nyatla.jp/nyartoolkit/wp/
[5]http://www.libspark.org/wiki/saqoosha/FLARToolKit/en

the characteristics of multi touch displays and b) integration of the physical world through tangible interfaces [6]. Different devices suit different interaction techniques. For example, a handheld tablet is very useful to play games, surf the web or read eBooks. In augmented reality, users usually manipulate data through a variety of real and virtual mechanisms and can interact with data through projective and hand-held displays. Tangible interfaces allow direct interaction with the physical world and virtual world using real, physical objects and tools. Tangible augmented reality (TAR) [10] combines the intuitiveness of tangible user interfaces (TUI) [28] with the abstractness of virtual objects. In a TAR environment, the user is normally in an *egocentric* view and is able to interact with virtual objects by using a TUI–based direct manipulation artifact. Common techniques within the TUI/AR are *paddle* [35], *cup* [36], or *box* shaped [41] props. 3D object selection and manipulation is possible by collision or proximity between the prop and a marker representing the 3D object. 3D object position and orientation is modified using a tilting, dropping, or hiding gesture using the prop (see Figure 2.3).

Natural Interaction in virtual environments is a key requirement for the virtual valida-tion of functional aspects in the design of PD rehabilitation programs. For example, in gait rehabilitation programs, patients are usually asked to pick up objects and per-form tasks with those objects. Natural interaction is the metaphor people encounter in reality: the direct manipulation of objects by using their hands (see Figure 2.4).

Moehring and Froehlich developed a glove–based pseudo–physical approach that tracks active markers to locate users' hands [48]. Their approach supports robust finger–based interaction of multiple hands, in which objects react to different hand and finger gestures. The main contribution is the pseudo grasping capabilities based on rigid body simulations (see Figure 2.4 (c)). A different glove–based natural inter-

Figure 2.3 Selection and manipulation using different prop techniques

a) Cube technique

b) ExPaddle technique

c) Paddle technique
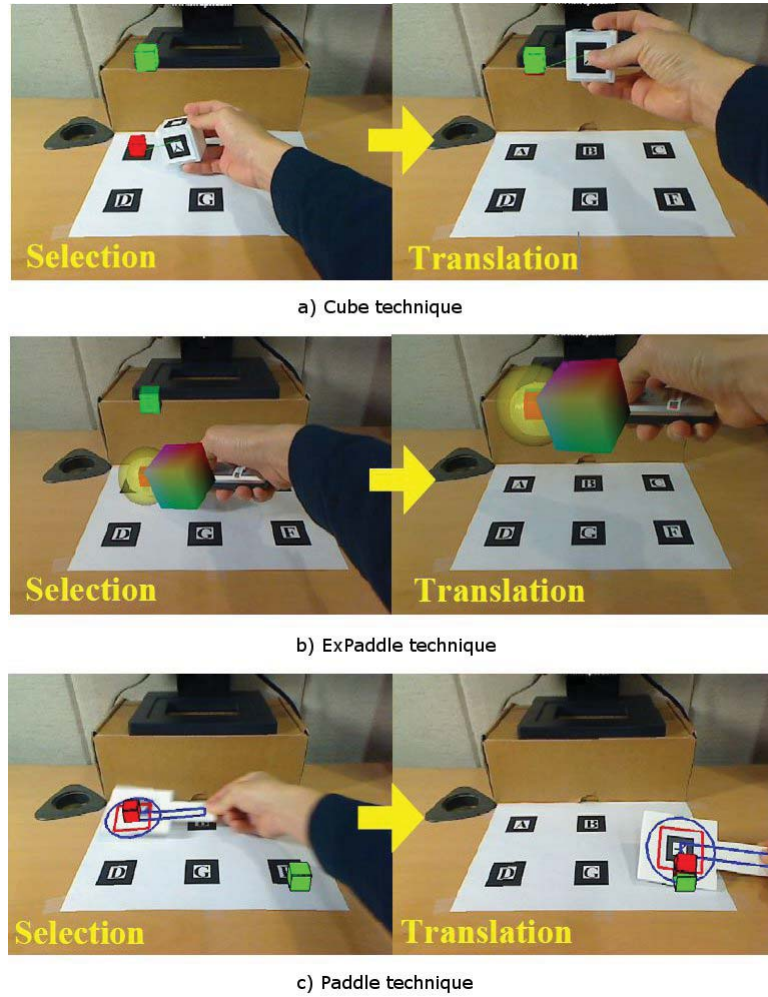
action approach was proposed by Thomas, W. Piekarski in [62]. The user operates an application with a device they called TinmithHand interface (see Figure 2.4 (b)). TinmithHand is a wearable approach that integrates head movement, hand tracking, pinch gloves, and a menu system to perform the different manipulation tasks. ARtoolkit [34] was used to compute the 3D transformation matrix of the markers placed on the pinch gloves. When using the system, users are able to select and manipulate virtual object. Additionally, users can create and combine 3D primitives. A very similar gesture–based direct manipulation approach is presented in [13]. Two very important contributions of Buchmann et al., was the use of OpenGL stencil buffer

Figure 2.4 Different glove based natural interaction techniques
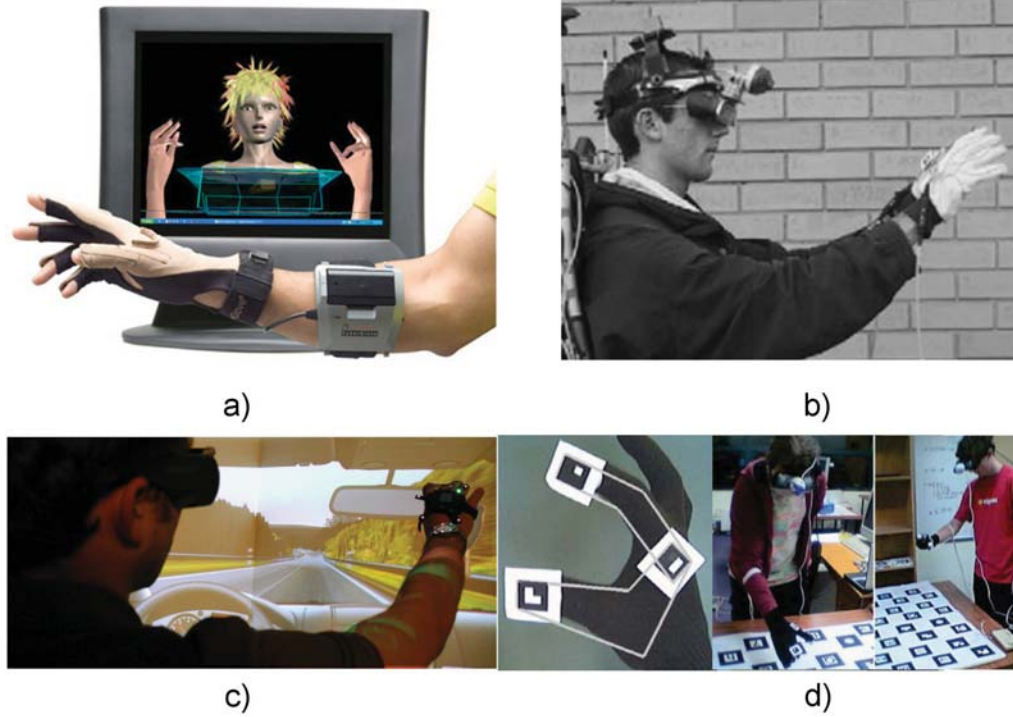
to render a segmented representation of the hands. Unlike previous AR interfaces, their device allows fingertip–based haptic feedback that enables users to "feel" virtual objects (see Figure 2.4 (d)).

As mentioned before, our system uses color–based skin classification to segment users' hands from the video signal to allow natural interaction. In our approach the segmented images are rendered directly on an OpenGL frame buffer object on top of the virtual information.

2.3   Navigation

The most intuitive way of navigation is natural walking. However, virtual environments still face various restrictions to allow unrestricted walking. One of the big issues of VEs has been the unfulfilled goal of enabling a person to move freely in the cyberspace without using metaphors which translate gestures to motion [65]. Most current setups do not offer the possibility of walking through VEs, or if they do, it is only in a very restrictive manner. In desktop–based metaphors, users simply navigate through the VE using keyboard, mouse or joystick, or similar input devices. This creates a sensory conflict, where the user is physically not moving, but receives visual input congruous with self–motion [57].

Innovative approaches to solve the navigation issue have emerged. Such approaches allow unencumbered movement within the virtual space through user self–motion. One example is the so called Gaiter System [59], which evaluates the movements of users to simulate motion without using a special floor or treadmills. However the real movement is limited by the room dimensions. The omnidirectional treadmill (ODT) [57] uses orthogonal belts which are made up of rolls. This machine facilitates omnidirectional unrestricted walking in the infinite virtual environment, within a finite real world footprint. A different approach, the Torus Treadmill [6], uses several belts which form a complete torus [29]. These advanced walking devices have usually been combined with Cave Automatic Virtual Environments (CAVE)[15] to maximize the immersive experience. The virtual theatre and the omnidirectional treadmill[6] are a commercially available setup that provides a full immersive experience.

---

[6]http://www.mseab.se/The-Virtual-Theatre.htm

Figure 2.5 Advanced walking systems
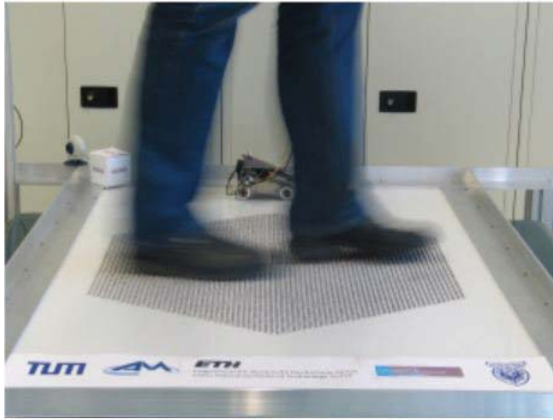
a) CyberCarpet

b) ODT and CAVE system

c) CyberWalk

d) TORUS treadmill

2.4 Skin segmentation using color pixel classification

In this section, we present a brief overview of the work that has been done by other research groups in an attempt to correctly classify skin color pixels. Existing skin–color classification approaches can be grouped into two basic categories: machine learning based approaches and approaches based on simply setting some sort of threshold on

color values [63]. Machine learning approaches have been demonstrated to achieve better results because pixels with colors corresponding to skin are clustered as non–linear distributions inside color spaces (see Figure 2.6). Such distributions are not separable by using simple thresholding [22]. One simple approach is to explicitly define a boundary. For example, in [12], a pixel is labeled as skin if the ratio between its red and green components falls between certain lower and upper boundaries that were obtained empirically.

Figure 2.6 Skin color distribution
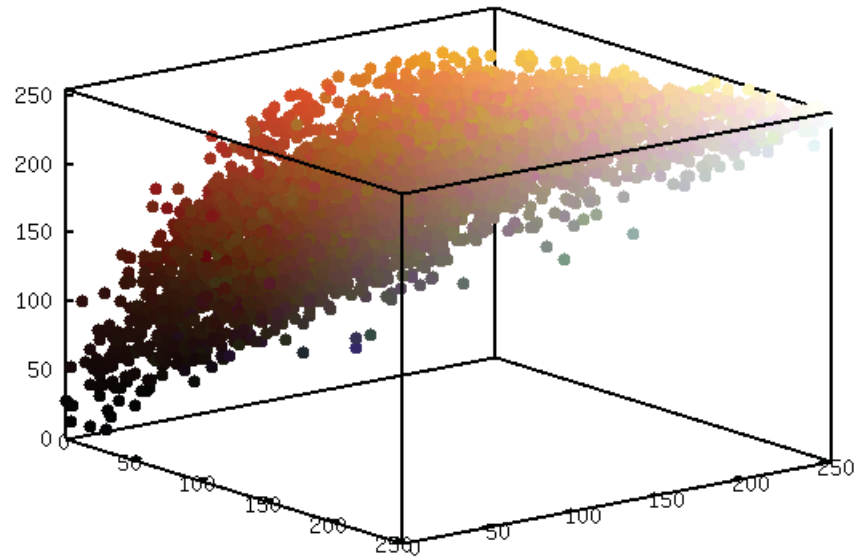
2.4.1   Machine learning approaches

Machine learning approaches can be subdivided into statistical approaches and physical–based approaches. Statistical approaches are most frequently used for skin classification. An example of physically–based approaches is [58], in which the authors made direct use of a physical model of skin reflectance. The reflectance model is used to

Figure 2.7 Non–skin color distribution
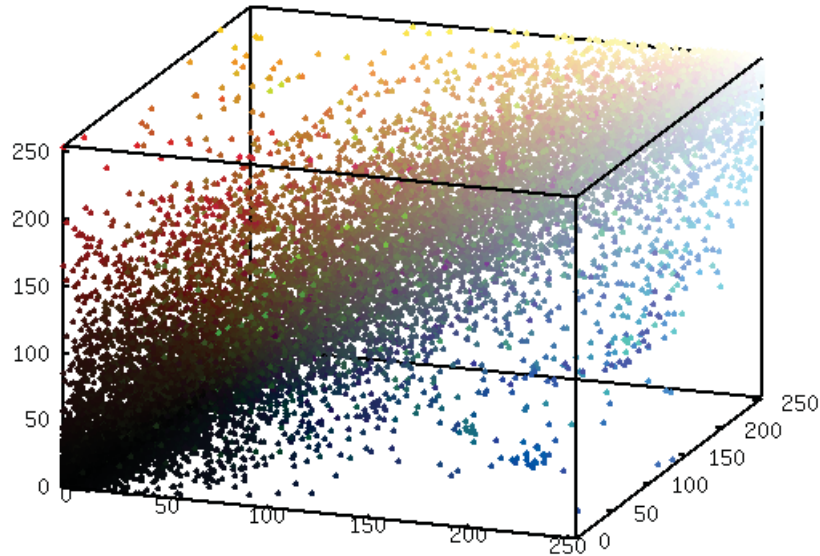
discount a known, time–varying illuminant to obtain color constancy. classification tends to be more accurate due to the algorithm use of strong prior knowledge such as camera and illumination parameters, as well as initial image segmentation. However, such information may not be readily available in analysis of everyday image sequences.

Statistical approaches can be subdivided further into parametric approaches and non-parametric approaches. Parametric statistical approaches represent the skin–color distribution in parametric form, such as a Gaussian model [16]. However, skin–color distribution is oftentimes multimodal and cannot be adequately represented as a single Gaussian in color space [60]. Mixtures of multiple Gaussians have been proposed and Expectation Minimization (EM) algorithms have been used to adapt models based on observed data. More advanced adaptive statistical approaches have emerged as an intent to deal with time varying illumination and its repercussions to skin appearance [55].

Nonparametric statistical approaches are a more general representation of the color distribution [52], where histograms are used to represent density in color space. A major drawback is that the histogram approach requires a considerable amount of training data. Different adaptations of a two–class Bayes classifier is by far the most used classifier among the literature. We believe that the reason is its non–iterative nature and its ability to learn the probability density function regardless of the complexity of the underlying distribution of skin–like color pixels.

In this thesis, we implemented a skin color classifier based on an artificial neural network approach. Also, an additional threshold–based color classifier was used to segment objects of interest such as baskets and a watering can.

2.5 Virtual environments in Parkinson's Disease research

Navigation can be seen as an interaction between mobility and an environment that requires the rapid integration of information from visuospatial input, kinematic input and memory. Navigation deficits involving visual processing have been reported in PD [68][17], and may contribute to gait impairment, increased risk of falls and inefficiency in completing tasks. Virtual Reality is a technology that has been used for assessing and rehabilitating such complex deficits. VR uses computer graphics software to create tridimensional virtual environments that visually immerse users, resulting in the perception that those environments are real. Virtual reality has been used in rehabilitation of gait and cognition in a variety of neurological conditions [42][11]. This technology has demonstrated efficacy for both assessment and treatment [69].

The field of virtual reality research in PD has grown rapidly in previous years. Many studies have utilized non–immersive systems that do not allow ambulation. Studies have focused on aspects of reaching, problem–solving and navigation using non–ambulatory, desktop–based systems [3][45][44]. Kaminsky et al. evaluated the effect of visual and auditory cues along with VR to simulate the real–world experience during ambulation [32]. Mirelman et al. used immersive virtual environments to provide visual context and cognitive/motor challenges in a VR gait–training program. However, the trajectory of ambulation was restricted to treadmill walking [46]. Hollman et al. [25], used a curve display and a treadmill to study whether or not gait instability is prevalent when people walk in immersive virtual environments. Their results suggest that the use of treadmills combined with VEs can cause instability in stride length and step width as well as variability in stride velocity.

In a previous in–home VR based project [33][64], the research group developed a fully simulated house that delivered visual information in the form of static contextual cues typical of a home environment such as furniture, doorways, walls, etc. In that study, the goal was to observe patients with PD ambulating freely without the inherent veering restrictions of a treadmill in a more "familiar" virtual environment. There were two main objectives in the study: the first was to assess the ability of individuals with PD to tolerate ambulating in an immersive simulated home environment. The second was to evaluate how deficits in the proposed environment affect the duration of task completion between two conditions requiring navigation in the virtual home: one utilizing a visual guide through the virtual home and the second without a visual guide. A head mounted display was worn by patients with PD to visually immerse them inside the three–dimensional virtual environment. Based on patients' orientation and ambulation in the real world, a third person was in charge of patients' navigation inside the virtual home using an experimenter–driven "Wiz-

ard of Oz" controlling scheme. The study was conducted with patients with PD and controls in a variety of navigation tasks such as line following tasks and free–form room–to–room navigation tasks. Results from that study were both interesting and valuable, indicating significant potential for the use of virtual worlds in creating ecologically valid research and rehabilitation environments for PD [33][64]. Nevertheless, the "Wizard of Oz" keyboard–based controlling scheme did not completely allow us to study neither navigation nor gait impairment, leaving doubts of whether or not virtual environments could be really used as auxiliary tools for gait–related rehabilitation programs.

Six Degrees of Freedom (6DOF) tracking devices have been used together with VR systems to monitor the position and orientation of selected body parts of users. When used on a head mounted display, the position and orientation of the head can be measured. This information defines the user's viewpoint in the virtual world and determines which part of the VE should be rendered to the visual display. The information delivered by tracking devices can be used to simulate navigation [7]. However, despite their huge cost, tracking devices are still prone to failure due to interference, out–of–range distances, or failure due to sensitivity to environmental factors. Depending on the technology used, 6DOF position trackers can be sensitive to large metal objects, various sounds, and objects coming between the source and the sensor.

2.6 Discussion

Previously published Parkinson's disease studies have focused on aspects of reaching, problem–solving and navigation. However, patients had to use Joysticks or the key-

board to both interact with the environment and navigate through it. In this thesis, instead of using desktop–based devices, commercial data gloves [54], or fiducial markers to track the position of user's hands [13], we segment the user's hands and make them visible inside the VE to allow natural interaction both with real and augmented objects.

Regarding navigation, researchers have evaluated the effect of visual and auditory cues delivered via virtual environments to enhance the real–world experience and cognitively challenge patients during ambulation using treadmills [46]. Many other non–Parkinson's Disease studies have utilized omnidirectional treadmills combined with CAVE systems to immerse people inside virtual environments. Such advanced configurations allow people to freely navigate in any direction inside the VE without restrictions. Unfortunately, the size, complexity, but above all, the price of such systems is so high, that makes it infeasible to use them for rehabilitation. The AR system we developed as part of this thesis, takes advantage of the vision–based tracking characteristics of Augmented Reality to obtain the 6DOF transformation of the camera. We use such transformation to emulate a head motion tracking system. The use of this 6DOF head tracking system allows patients with PD to freely navigate inside virtual environments without using any kind of treadmill or inertial/hybrid tracking devices.

Chapter 3

System Design and Development

3.1 Overview

In Chapter 2 we discussed how immersive virtual environments have been used for training and rehabilitation in many different areas. Virtual environments have been particularly useful to simulate tasks that take place in dangerous places. In addition, virtual environments help minimize costs because they allow to model places that would otherwise be extremely expensive or difficult to reproduce for training purposes.

A significant amount of money has been invested in areas such as the military and the oil industry to build advanced virtual environments. These environments rely on specialized equipment that allows users to navigate the virtual world. Examples of equipment used in these systems include omnidirectional treadmills for navigation, hybrid tracking devices for pose estimation, CAVE systems for visualization and data gloves for interaction. These devices are intended to provide a complete sense of

immersion.

The problem with technologies such as virtual reality, is that they require the use of expensive devices in order to be able to provide a sense of immersion to the user. Therefore, virtual reality systems can prove difficult to implement if financial resources are limited.

The main objective of this thesis is to develop a system that can provide the user with a sense of immersion without requiring expensive equipment. In addition, some technologies are impractical for this particular application. They might not be suitable for in–clinic use, equipment might be bulky, heavy, or awkward for patients to use (especially seniors or people who have mobility issues; for example, treadmill–based systems are not suitable for them). This is important, because our system is intended to be portable and transferable so that it can be used in any hospital without requiring a huge investment. Our system will allow physicians to observe patients with PD as they perform daily life activities in the virtual environment.

In an attempt to reduce costs, some Parkinson's disease research groups have developed systems that use desktop–based metaphors, which employ simpler devices such as off–the–shelf game controllers for interaction. In those systems, navigation is implemented through the use of common treadmills. That approach, however, has not given satisfactory results. One of the problems is that a common treadmill provides limited range of movement since it only moves in one direction. This makes it impossible to reproduce real–life activities under context, which is very important for a successful rehabilitation.

Our wearable augmented reality system is transcendental and innovative because it provides natural interaction and free navigation. This is made possible by our

implementation of a skin classification algorithm that uses artificial neural networks to identify the user's hands and overlay them over the virtual environment. In terms of navigation, the only limitation of our approach would be the physical space available.

3.2 The wearable AR framework

In this section, we describe the three main components of our wearable AR framework. In Section 3.2.1, we describe the hardware requirements of our system as well as the particular hardware selections made for our current implementation. In Section 3.2.2, we describe the physical setup we used to perform our experiments. Finally, in Section 3.2.3, we describe the software architecture of our system.

3.2.1 Hardware

Our approach needs a camera system to sense the environment and provide a source video stream for augmentation and positioning/orienting the user, a computer to run our software and do all the processing involved to construct, compose, render, and produce the environment as the user should see it, and a head–mounted display for presenting the virtual environments to the user. The main aspects that we considered to decide the hardware to be used in our framework were: weight, computing power and connectivity.

The laptop computer. One of the main hardware components in our framework is a laptop computer. We chose one that was light so that it could be fit into a small backpack. This is important, because it is what makes our system wearable. Our

objective was to minimize the patients' awareness regarding the fact that they are carrying or "wearing" a laptop. We consider this to be crucial to provide a better sense of presence, since the patient can concentrate on the task at hand without worrying about the laptop. Another important factor in our decision was computing power since our system renders 3D graphics and processes video at the same time. In addition, we needed a laptop with support for an Internet wireless connection, video output and USB ports. We chose the ASUS UX31 because it was the lightest Windows–based computer that complied with our requirements. See Figure 3.1 (a).

Figure 3.1 The three devices used in our framework

a) ASUS UX31   b) VUZIX iWear 920VR   c) VUZIX iWear CamAR

The head mounted display. This device is vital in our system, because it is through it that the user sees the virtual environment in first–person (i.e. as if patients were using their own eyes). Figure 3.1 (b) shows the VUZIX iWear 920VR HMD we are using. This model is light and supports a resolution of up to 1024×768 pixels.

The camera. This device is used to capture video at 30Hz. The video is processed by computer vision algorithms in order to compute the position of the camera relative to the real world. In Section 3.2.3, we describe the computer vision algorithms in more detail. We decided to use a Dinex CamAR webcam, which is shown in Figure 3.1 (c). This model is designed so that it can be easily attached to the VUZIX iWear 920VR

HMD.

3.2.2 Setup of the physical space

In order to use our system, a physical space is required in order to install the fiducial markers needed to represent the virtual world. In this section we will describe the physical setup of the space we used for our experiments.

In order to setup our system, the London Health Sciences Center provided us with a room that measures $6.68\ m \times 4.92\ m$. The room is enclosed by four vinyl walls over which we mounted fiducial markers. We also installed fiducial markers in the floor. *Fiducial markers* are points of reference that a computer vision system uses to measure the position of the camera with respect to each fiducial marker. The fiducial markers in our system are unique black and white patterns printed on a material known as *coroplast*. Black and white fiducial markers are easier to detect because they provide high contrast. Figure 3.2 shows a photograph of the physical space with the fiducial markers. As we can observe, we installed fiducial markers of different sizes. Bigger markers are used to track the position of the camera from long distances. Smaller markers are used so that the user can interact with virtual objects from shorter distances. We used five different marker sizes: $45 \times 45\ cm$, $30 \times 30\ cm$, $20 \times 20\ cm$, $15 \times 15\ cm$ and $10 \times 10\ cm$. We installed and configured 110 fiducial markers in total.

As we can observe in Figure 3.2, the biggest markers, which measure $45 \times 45\ cm$, were installed in the bottom and top of the walls. The reason for this is that both the top and bottom of the walls are farther with respect to the point of view of the user. We can also observe that the markers are smaller in size as they approach the

Figure 3.2 Picture showing the physical space in which installed fiducial markers in the walls and on the floor.

level that corresponds to a person's eyes when looking straight ahead (approximately 1.70 m). In order to compute the position of the camera with respect to the markers in the room, the system must know the 3D position of each marker with respect to a *specific point of reference* in the real world. Therefore, we measured the 3D position of each marker with respect to the point of reference. Figure 3.3 shows a top view in a 2D coordinate system, where the positive x axis represents the width of the physical space and the positive y axis corresponds to the length of the physical space. We can observe that we represented the physical space in the upper right quadrant in the Cartesian plane coordinate system, so that all points have positive values. The shaded area represents the physical space area on quadrant I, which goes from (0,0)

Figure 3.3 A 2D, top view representation of the physical space where we conducted our experiments.
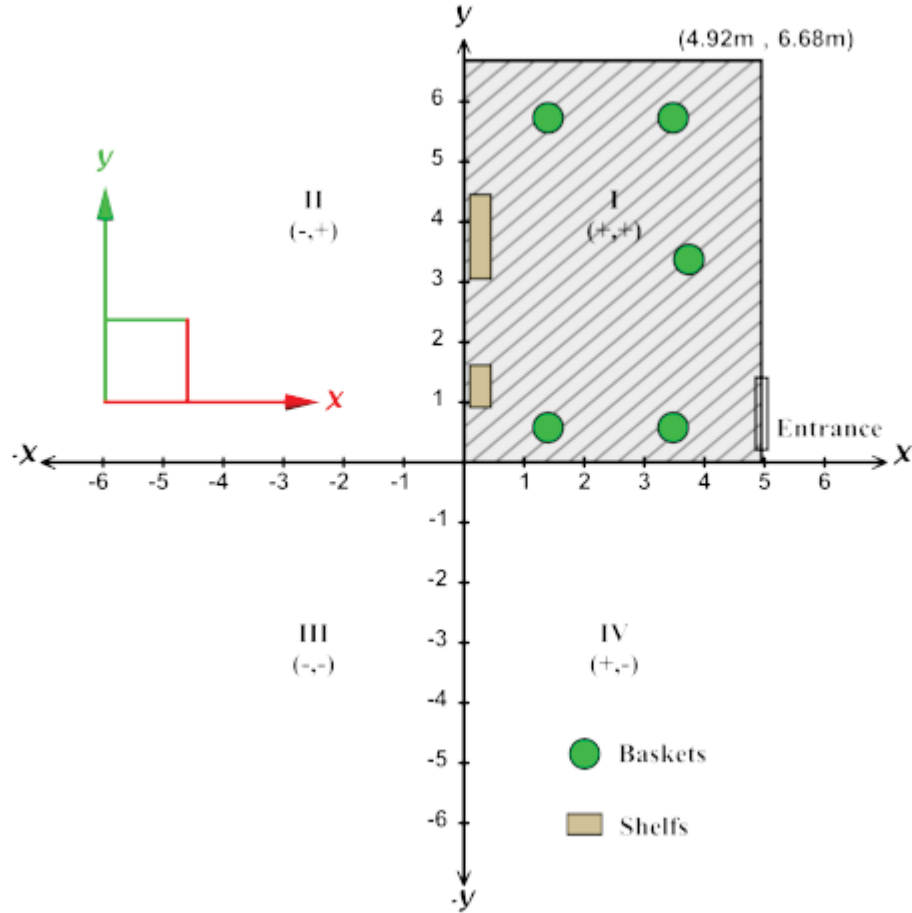
to (4.92,6.68). Our *point of reference* corresponds to the origin (0,0) in Figure 3.3.

Figure 3.4 shows a 3D representation of the physical space. In this representation, the width of the room is represented by the positive x axis. The positive y axis represents the length of the room. The height of the room is represented by the positive z axis. Each fiducial marker has a representation in a 3D space, which we use to measure and extract the 3D coordinates of each one of them. These coordinates are used by the AR Driver module described in Section 3.2.3. Figure 3.5 illustrates the 3D

Figure 3.4 A 3D representation of the physical space we used to conduct our experiments.
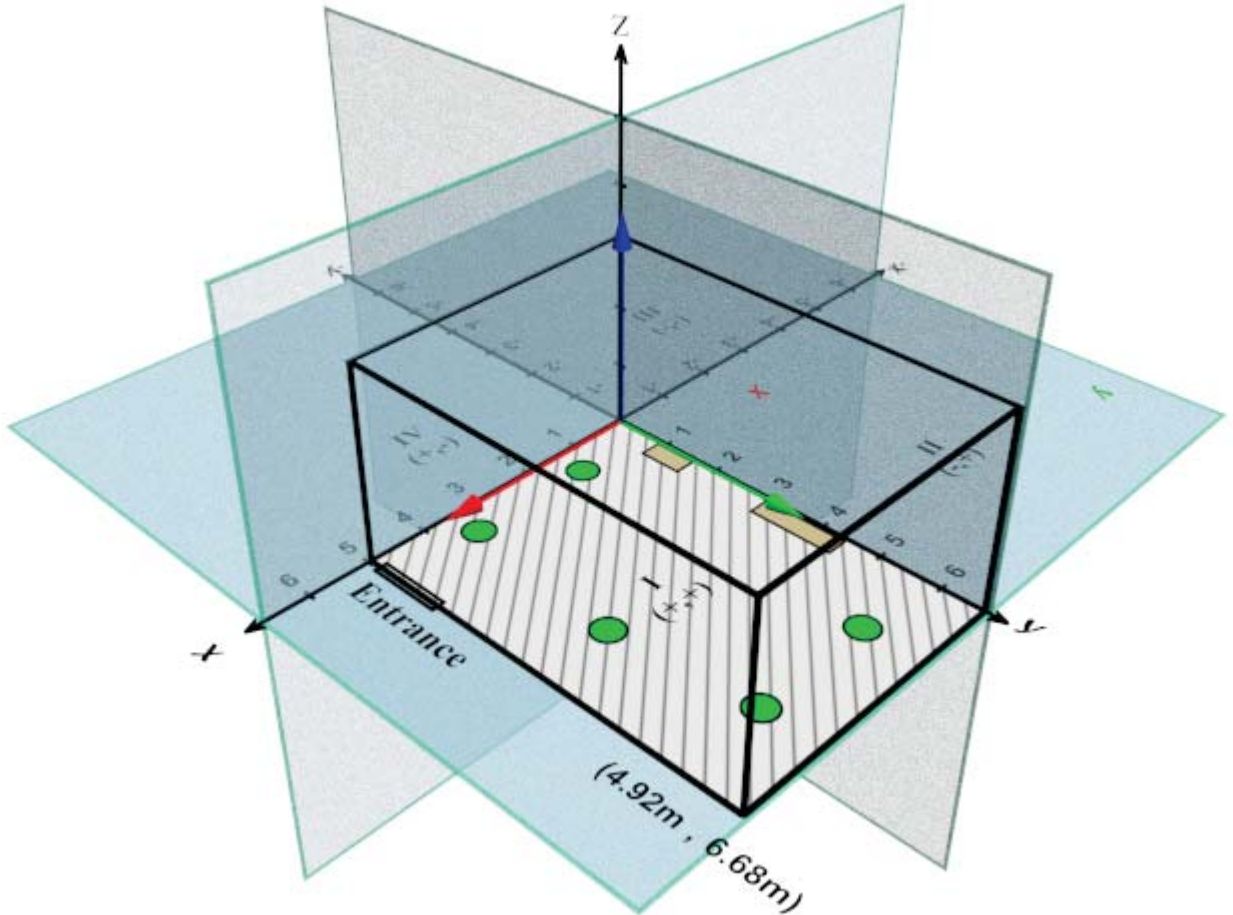
representation of our fiducial markers. In Chapter 4, we will describe the activities that were carried out by patients in our experiments.

3.2.3 Software

In this section we describe the software component of our AR framework. This software is novel because no other Parkinson's disease research group has used augmented reality to create immersive virtual environments. Even though our approach to allow

Figure 3.5 A 3D representation of the physical space where we conducted our experiments, which shows the actual position of the fiducial markers.
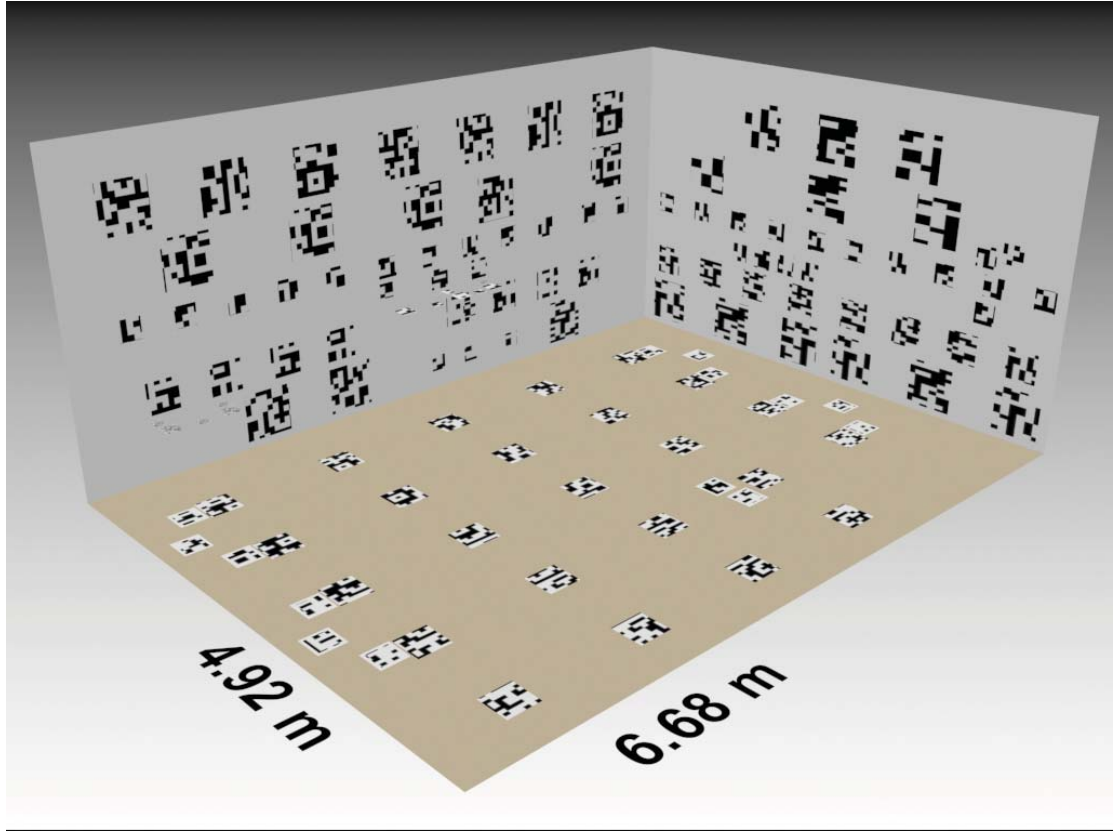
natural interaction is simple, users can select and manipulate real and *augmented* objects without the need for external devices. Another innovative aspect of our software is that it uses a real–time skin classifier based on artificial neural networks. The system was developed so that it would be easy–to–use and intuitive because it is intended to be used by patients with Parkinson's disease.

Our system is composed of four main modules: `CoreSystem`, `VideoSource`, `ARDriver` and `ScenarioManager`. Figure 3.6 shows the architecture of our system and illustrates how these four modules interact. As we can observe, the `VideoSource` module captures and processes the video signal. The `ARDriver` module computes the transformation matrices of the 3D objects. These matrices are fed to the `ScenarioManager`,

which renders the final scene. Each of these four modules will be discussed in the following sections.

Figure 3.6 System architecture
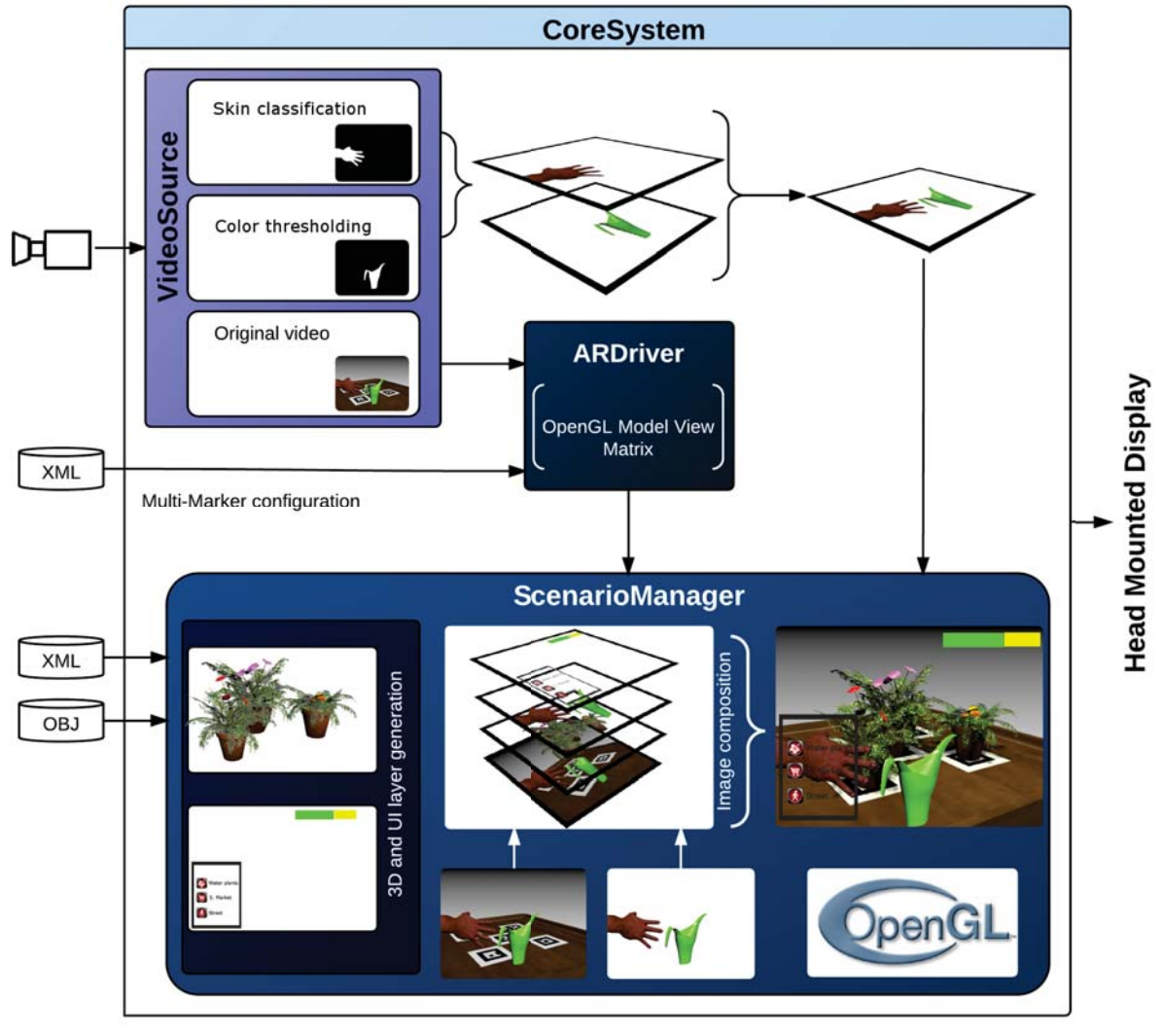
3.2.3.1 Core system

The `CoreSystem` module manages the data structures that are used by other modules. In addition, the `VideoSource`, `ARDriver` and `ScenarioManager` modules are

instantiated from `CoreSystem`. The `CoreSystem` module also manages the GUI and user actions in general. One of the main advantages of our system is that it allows creating and administering multiple scenarios without modifying or recompiling the source code. `CoreSystem` uses two XML configuration files in order to manage the structure and behavior of the GUI and 3D scenarios. Therefore, to create a new scenario, the operator of the system only needs to add new entries in two of the configuration XML files. We discuss these two types of configuration files below.

GUI and scenario configuration file. This XML file contains the information needed to display the GUI and to categorize the panels from which the different scenarios can be loaded. Figure 3.7 shows an example of a GUI configuration file. The system operator selects a scenario by accessing a panel that shows different *categories*. For example, in our experiments we had three categories: "Super market", "Watering the Plants" and "Street Walk". Lines 3, 6 and 9 in Figure 3.7 show the definition of these three categories. Figure 3.8 shows a screen capture of how these categories are displayed in the GUI. Each category is composed of one or more scenarios that represent different tasks to be performed by the patients. These tasks will be discussed in Chapter 4. Lines 10, 11 and 12 show the "5 seconds", "10 seconds" and "15 seconds" scenarios that correspond to the "Street Walk" category in line 9 of Figure 3.7. When the system operator selects a scenario, `CoreSystem` creates the necessary data structures using the information contained in the `xmlconfig` attribute. For example, if the operator selected the "10 seconds" scenario, `CoreSystem` would open the `trial_10.xml` file to obtain information on the behavior of the 3D models that correspond to that scenario.

Figure 3.7 GUI and scenario configuration file

```
 1 <?xml version="1.0"?>
 2 <ARSTUDY_Scenario_DB version="1.0">
 3     <Category name="Super Market" image="Super_Market/A.png">
 4         <Scenario name="Trial One" image="cone.png"
              dirCFG="Super_Market/Trial_One" dirModels="Super_Market/models"
              xmlConfig="trial_one.xml" />
 5     </Category>
 6     <Category name="Watering Plants" image="Watering_Plants/B.png">
 7         <Scenario name="Trial One" image="cone.png"
              dirCFG="Watering_Plants/Trial_One"
              dirModels="Watering_Plants/models" xmlConfig="trial_one.xml" />
 8     </Category>
 9     <Category name="Street Walk" image="Street_Walk/C.png">
10         <Scenario name="5 Seconds" image="tl_5.png"
              dirCFG="Street_Walk/Trial_One" dirModels="Street_Walk/models"
              xmlConfig="trial_5.xml" />
11         <Scenario name="10 Seconds" image="tl_10.png"
              dirCFG="Street_Walk/Trial_One" dirModels="Street_Walk/models"
              xmlConfig="trial_10.xml" />
12         <Scenario name="15 Seconds" image="tl_15.png"
              dirCFG="Street_Walk/Trial_One" dirModels="Street_Walk/models"
              xmlConfig="trial_15.xml" />
13     </Category>
14 </ARSTUDY_Scenario_DB>
```

3D models and behavior configuration file. This file is used to configure how a scenario is to be rendered. It also defines the 3D models that will be used to create the virtual environment. In addition, it defines which fiducial markers will be used in order to compute the 3D transformations for each object. Figure 3.9 shows an example of a 3D Model and Behavior XML configuration file. Line 5 defines a 3D object with the name "`t1.obj`" of type "`AnimatedTexture`". In our system we defined three types of 3D objects:

- **Scenography objects** are used to render the scenography of the virtual environment. In this type of objects, several fiducial marker configuration files are associated with the 3D model. The different configuration files are organized according to a hierarchy that dictates which fiducial markers have higher priority during the tracking process.

Figure 3.8 Graphical User Interface showing the main menu and sub–menus to select the different scenarios

a) Graphical User Interface: category selection   b) User Interface: Scenario selection under street walk category
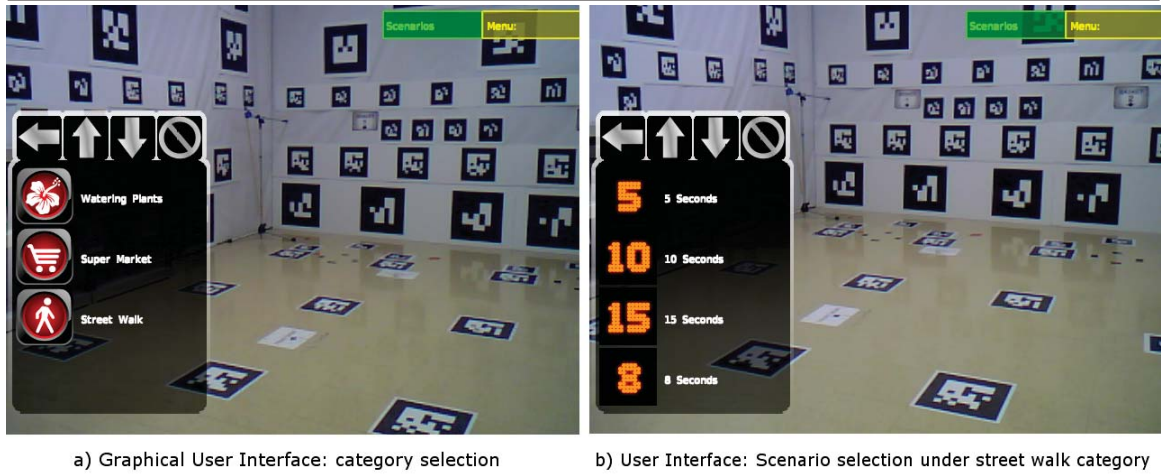
`CoreSystem` creates a data structure which contains the fiducial markers hierarchy, that is later used by `ARDriver` module to achieve better tracking accuracy.

- **Animated texture objects** are 3D objects with a behavior that consists of changing their texture dynamically. For example, this type of object was used to represent a pedestrian traffic light with 9 transitions (see figure 3.9, line 5).

- **Interactive objects** can be selected and manipulated by users. For example, in one of the scenarios we ask users to pick up different *augmented* cereal boxes. Each cereal box is defined as an interactive object.

3.2.3.2 Video source

This module is one of the most important components of our system because it captures the video signal and detects/segments objects of interest such as the hands of the patient. This module is divided into three main functionalities: **Video capture**, **Color thresholding** and **Color based skin classification**. We describe these three

Figure 3.9 3D models and behavior configuration file

```
1  <?xml version="1.0"?>
2  <ARSTUDY_ModelConfig_DB version="1.0">
3      <Scenario name="Trial One">
4          <ModelConfiguration objModelFile="room.obj"
              type="Scenography" xmlMarkerConfig="room45_5.xml"
              xmlMarkerConfig2="room30.xml"
              xmlMarkerConfig3="room20.xml"
              xmlMarkerConfig4="room20_5.xml"
              xmlMarkerConfig5="room15_5.xml"
              xmlMarkerConfig6="roomFloor.xml"
              xmlMarkerConfig7="room10_5.xml"/>
5          <ModelConfiguration objModelFile="tl.obj"
              type="AnimatedTexture"
              xmlMarkerConfig="markerData151.xml"
              xmlMarkerConfig2="markerData151.xml" transitions="9" />
6      </Scenario>
7  </ARSTUDY_ModelConfig_DB>
```

aspects below.

Video capture. The captured video signal is sent to the `CoreSystem` module so that the `ScenarioManager` module can incorporate the original video signal as background over which the 3D objects are rendered.

Color thresholding. The `VideoSource` module segments objects of interest using a simple thresholding technique to classify green objects. This classifier generates a black–and–white image that is used in combination with the results of the skin classifier to generate a mask which is used by the `ScenarioManager` module.

Color based skin classification This is the most important functionality of `VideoSource` because is the feature that allows natural interaction. Due to the skin classifier importance, in this section we describe the entire process we followed to train the Artificial Neural Network (ANN) classifier.

An artificial neural network can be defined as a machine learning algorithm which is inspired by the structural and functional aspects of biological neural networks. A neural network consists of an interconnected group of artificial neurons [47]. We considered neural networks to be a good choice for our skin classifier, because among the other options we tested (Support Vector Machines and simple thresholding), ANN gave us the best experimental results.

Processing the training database   Jones et al. developed a labeled skin classification dataset known as the Compaq dataset[31]. To build the dataset, researchers at Compaq collected almost 13,640 images from the Web and manually labeled each pixel to generate ground truth images. They used Portable Bitmap (PBM) files to store a ground truth image for each file. Each PBM file contains two types of labels: pixels corresponding to skin are labeled as 1 (white) and any other pixel is labeled as 0 (black). Figure 3.10 (b) shows an example of a labeled mask. Figure 3.10 (a) shows the original unlabelled picture.

Dataset refinement   We processed all the 13,640 images contained in the Compaq dataset by extracting the red, green and blue components of each pixel. A *OpenCV* application was developed to generate the training files. Each row in our training file represents a pixel and there are four columns per row. In the first column, we used the ground truth label corresponding to the pixel. In the remaining three columns, we have the red, green and blue components of that pixel. An example of how the training examples are stored is presented below.

```
1 255 239 213
```

```
0 106 090 205
```

Figure 3.10 Image classification process

a) Original image   b) ANN classification result   c) Segmentation process
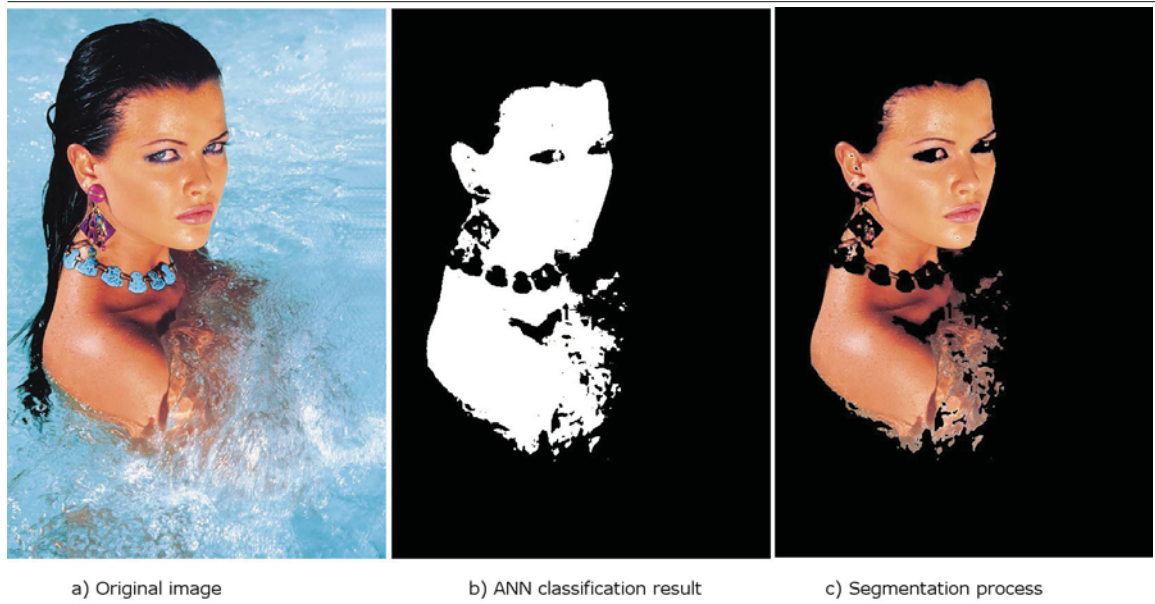
In order to better understand the dataset, we plot the skin color distribution. Figure 2.7 shows the color distribution for non–skin pixels whereas Figure 2.6 shows the color distribution for skin–pixels. In Figure 2.6, it is interesting to observe that the color distribution is very compact, since it is clustered mainly around the diagonal gray line [31] in the RGB color space. By looking at both figures: 2.6 and 2.7, it is easy to observe that separating skin–pixels from non–skin pixels is non–linearly classifiable.

Cross–Validation In k–fold cross–validation, the dataset D is randomly split into k mutually exclusive subsets. These subsets are known as **folds**. These folds, represented by the expression $\mathfrak{D}_1, \mathfrak{D}_2, \ldots \mathfrak{D}_k$ are of approximately equal size. The k training sets are trained and tested k times. Each time $t \in \{1, 2, \ldots, k\}$, each of the k subsamples is used exactly once as the validation data. The k results from the folds can then be averaged (or otherwise combined) to produce a single estimation. In our experiments, we chose to use $k = 10$ because it has proven to produce good results and the trade–off between time and accuracy is good [38]. The cross–validation

estimate of accuracy is the overall number of correct classifications, divided by the number of instances in the dataset. Formally, let $D_{(i)}$ be the test set that includes instance $x_i = \langle v_i, y_i \rangle$, then the cross–validation estimate of accuracy is given by:

$$acc_{cv} = \frac{1}{n} \sum_{(v_i, y_i) \in \mathfrak{D}} \delta(\mathfrak{I}(\mathfrak{D}_{(i)}, v_i), y_i) \tag{3.1}$$

We used k–fold cross–validation to identify the optimal number of iterations for our ANN, following [47], we chose the stop condition for the final classifier where the accuracy over all k test was the highest.

ANN implementation overview We implemented the backpropagation algorithm in Octave. GNU Octave is a high–level interpreted language, mainly intended for numerical computations [19]. It provides good capabilities for matrix operations and optimization problems. In our implementation, the neuron model is the non–linear logistic function. Equation 3.2 shows the definition of the logistic function.

$$g(z) = \frac{1}{1 + e^{-z}} \tag{3.2}$$

The logistic unit has the nice property of being a function whose output is a non-linear function of its inputs, but whose output is also a differentiable function of its inputs [47]. This is useful in our context, because skin–colored pixels are not linearly separable from non–skin pixels. The cost function for our neural network is defined in equation 3.3.

$$J(\Theta) = -\frac{1}{m} \left[\sum_{i=1}^{m} \sum_{k=1}^{K} y_k^{(i)} \log(h\Theta(x^{(i)}))k + (1 - y_k^{(i)}) \log(1 - (h\theta(x^{(i)}))k) \right] + R \tag{3.3}$$

The second term of equation 3.3 (R) represents regularization. Regularization is a method that prevents overfitting and it works by penalizing complex models[47]. This regularization term is defined in equation3.4.

$$R = \frac{\lambda}{2m} \sum_{l=1}^{L-1} \sum_{i=1}^{s_l} \sum_{j=1}^{s_l+1} (\Theta_{ji}^{(l)})^2 \tag{3.4}$$

How much regularization is applied is determined by the magnitude of the parameter λ (lambda). With a small lambda, complex models are slightly penalized. In contrast, a large lambda causes more penalties to complex models. The cost function was minimized by using an Octave advanced minimization solver called `fminunc`.

Neural Network Architecture We trained several neural networks with three layers and the following characteristics:

- **Input layer**, consisting of three input units (red, green and blue components of a pixel, respectively)

- **Hidden layer**, consisting from 3 to 25 hidden units.

- **Output layer**, with one output unit

Figure 3.11 represents the structure of the neural networks we trained. Blue units represent our input layer, green units represent our hidden layer and the red unit, represents our output layer.

Parameter search It has been suggested that the number of hidden units are a function of the number of input and output units [9]. However, the actual number of hidden units is a complex combination of factors such as: target function complexity,

Figure 3.11 Artificial neural network architecture
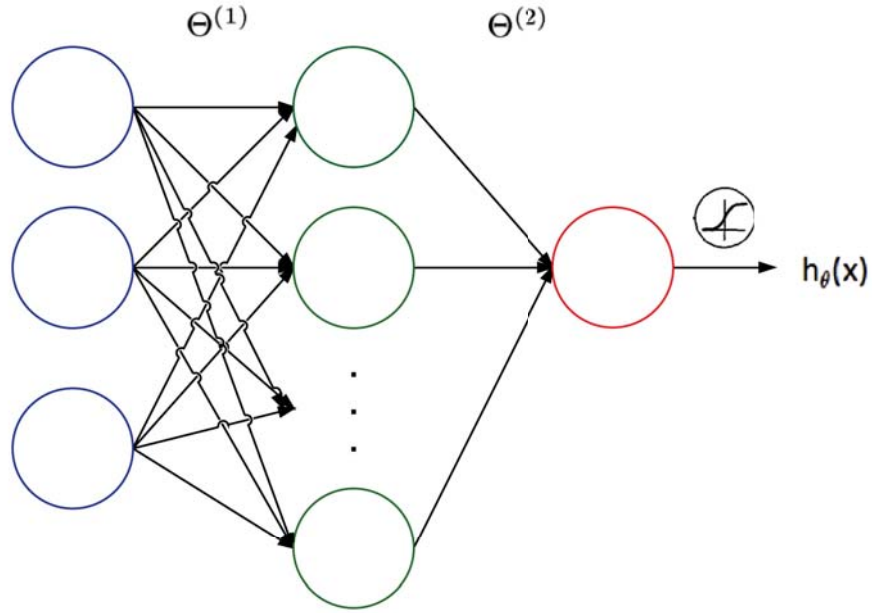
number of training examples, amount of noise in the data, number of inputs and outputs, etc. [53]. In order to find the neural network that worked best for the problem, several parameters were used for experimentation.

- **Number of iterations**, when minimizing the cost function, we tested up to 200 iterations.

- **Number of hidden units**, we tested between 3 and 25 hidden units.

- **Regularization parameter** λ, we tested with $\lambda = 0.05$, $\lambda = 0.5$, $\lambda = 1$, $\lambda = 2$, $\lambda = 5$ and $\lambda = 10$.

To test the robustness of the classifier, the neural networks were trained with two randomly generated training sets.

- The first training set contained 20,545 training examples.

- The second training set was composed by 78,405 training examples.

ANN results and application As mentioned before, in order to find the best possible parameters of our neural networks, we tested different sets of parameters. In our experiments, the parameters that performed the best were: 4 hidden units, $\lambda = 0.05$ and stopping at 68 iterations. We obtained an accuracy of 85% over the training set and the test set (see Figure 3.12). The ANN prediction model was

Figure 3.12 ANN with 4 hidden units and $\lambda = 0.05$
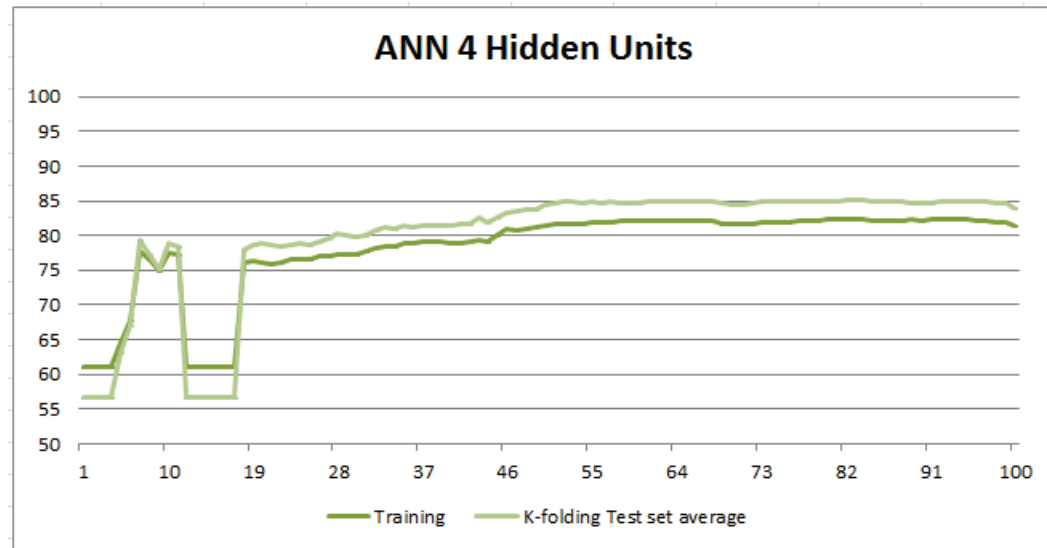
implemented in *OpenCV* and is part of the *VideoSource* module. The prediction is used every frame to segment the image and generate the "skin alpha channel mask". The skin mask is combined with a previous mask that contains the pixels segmented by the threshold–based color classifier. The resulting mask is attached as alpha channel of a new layer. Every frame, the *ScenarioManager* method blends the layer that contains the hands and objects of interest over the virtual environments.

Figure 3.13 Skin detection procedure
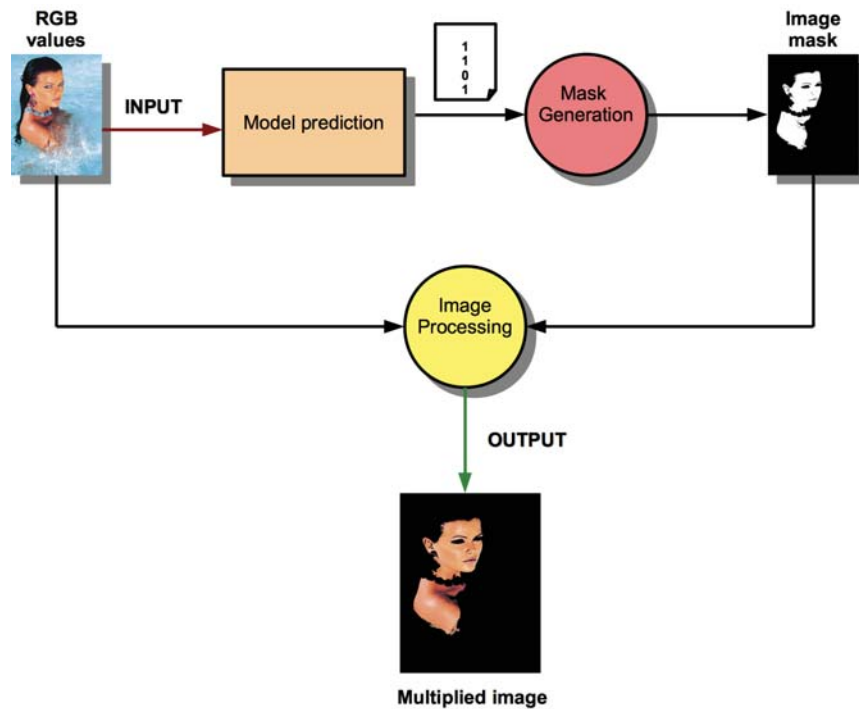
Figure 3.13 exemplifies the image prediction and composition. The RGB values from an image are fed to the model. The model predicts which pixels correspond to skin. Then, the *mask generation* process generates a mask in which skin pixels are labeled with 1 (white) and non–skin pixels are labeled with 0 (black). An image processing method multiplies the original image and the mask to obtain a new 4 channel image that contains alpha values. The multiplied image shows skin pixels only, pixels that do not correspond to skin have alpha values of 0 (transparent pixels).

3.2.3.3   ARDriver module

This module detects and extracts the position of the fiducial markers with respect to the camera. `ARDriver` receives an instance of the video signal from the `CoreSystem`

module and detects all of the fiducial markers on the current frame. Each 3D object is associated with a series of different markers. This is known was *multi–marker configuration.* Multi–markers are detected according to the hierarchy defined by the `CoreSystem` module (see Section 3.2.3). This hierarchy groups fiducial markers according to size. For example, a given multi–marker might be formed exclusively by four $10 \times 10 \ cm$ fiducial markers.

In our system, the multi–marker configurations are defined in an XML file, which follows the format defined by the augmented reality library known as ALVAR. Figure 3.14 is an example of the format used to define the offsets (expressed as 3D positions) of the corners of four fiducial markers with respect to our point of reference. ALVAR uses this configuration to compute the 3D transformation of each 3D object and the result is translated into the format required by the `ScenarioManager` module.

3.2.3.4 ScenarioManager module

This module receives and integrates information from `VideoSource` and `ARDriver`, in order to render the final scenario. Essentially, it integrates the video signal, the 3D models and the segmented objects of interest to create the augmented reality environment that the user perceives. In Figure 3.6, in the `ScenarioManager` box, we can observe an illustration of how these elements are merged. Additionally, `ScenarioManager` renders the GUI when necessary. The `ScenarioManager` renders the scene by performing the following actions:

1. `ScenarioManager` receives an instance of the original video feed from `VideoSource` and composes an *initial layer* over which the 3D models will be rendered.

Figure 3.14 Multi–marker configuration file

```
 1  <?xml version="1.0" encoding="UTF-8" standalone="no"?>
 2  <multimarker markers="4">
 3      <!--***********************************************************
 4         AR STUDY ROOM MULTIMARKER CONFIGURATION
 5         ***********************************************************-->
 6      <marker index="123" status="1">
 7          <corner x="34.00000o" y="343.600000" z="98.500000" />
 8          <corner x="34.00000" y="353.600000" z="98.500000" />
 9          <corner x="24.00000" y="353.600000" z="98.500000" />
10          <corner x="24.00000" y="343.600000" z="98.500000" />
11      </marker>
12      <marker index="124" status="1">
13          <corner x="12.6000" y="374.100000" z="98.500000" />
14          <corner x="12.6000" y="384.100000" z="98.500000" />
15          <corner x="2.60000" y="384.100000" z="98.500000" />
16          <corner x="2.60000" y="374.100000" z="98.500000" />
17      </marker>
18      <marker index="125" status="1">
19          <corner x="12.2000" y="419.300000" z="98.500000" />
20          <corner x="12.2000" y="429.300000" z="98.500000" />
21          <corner x="2.20000" y="429.300000" z="98.500000" />
22          <corner x="2.20000" y="419.300000" z="98.500000" />
23      </marker>
24      <marker index="126" status="1">
25          <corner x="33.10000" y="397.600000" z="98.500000" />
26          <corner x="33.10000" y="407.600000" z="98.500000" />
27          <corner x="23.10000" y="407.600000" z="98.500000" />
28          <corner x="23.10000" y="397.600000" z="98.500000" />
29      </marker>
30  </multimarker>
```

2. ScenarioManager receives from CoreSystem the list of all the 3D models that need to be rendered. Those models are read from disk once and loaded into memory using OpenGL Display Lists. A display list is a group of OpenGL instructions that are compiled once. These instructions contain vertex, pixel and texture data that are copied to the video memory. The advantage is that they can be used repeatedly without re–evaluating and re–transmitting data over and over again to draw each frame.

Figure 3.15 3D scenarios with baked lighting

a) Grocery store   a) Living room

3. `ScenarioManager` associates the `OpenGL Display Lists` stored in memory with the transformation matrices received from the `ARDriver` module, to correctly project the 3D models into a *second layer*.

4. `ScenarioManager` receives an image that contains the segmented user hands and objects of interest from the `VideoSource` module. With this information, it creates a *third layer*.

5. Finally, `ScenarioManager` merges the three layers mentioned above to generate the scene by using the `OpenGL Alpha Blending` capability.

To make scenarios more realistic and to enhance depth perception, our 3D models contain lighting information. In this way, the 3D objects seem more real since they cast shadows. Figure 3.15 shows two of our scenes. We can observe that objects cast shadows, making the scene more realistic.

Chapter 4

Experiments Protocol

One of the objectives of our thesis was to evaluate whether augmented reality can be used as a support tool in the development of rehabilitation programs for patients with Parkinson's disease. This thesis is part of a pilot study whose objective is to explore if augmented reality can be used as an alternative to create immersive virtual environments. The short–term objective of that pilot study is to assess whether or not patients react in a similar way when performing tasks in a virtual environment and the real world. The long–term objective of that pilot study is to develop different context–based task–oriented rehabilitation programs. This thesis is part of the ongoing work that is being done to fulfill those objectives. The results from our thesis will contribute to the pilot study by comparing the performance of patients when doing tasks in a virtual environment and the real world. To perform this comparison, we carried out a series of experiments based on a protocol that was developed by the Movement Disorders Centre which is headed by Dr. Mandar Jog. The protocol was approved by the Human Subjects Research Ethics Board at the University of Western Ontario. We used this protocol as a case study because we are interested in study-

ing if augmented reality is suitable for the development of rehabilitation or training applications. To evaluate our system, we performed a series of experiments that are designed to challenge patients in a similar way as it is done in regular rehabilitation programs. We used our system to observe how patients respond to cognitive, motor and executive–function challenges.

For our experiments, eleven participants between the ages 50 and 80 were recruited using a convenience sampling technique from the Movement Disorders Centre at London Health Sciences Centre. nine of these individuals had Parkinson's disease (patients with PD), while two of them did not (controls). The criteria for inclusion and exclusion of participants in the trials were determined by the Movement Disorders Program. For example, patients with a high–level of dementia were excluded from the study because they are unable to follow instructions. Patients that present any type of *freezing of gait* were excluded because they are prone to falling. The experiments were performed at the London Health Sciences Centre South Street Hospital. The procedures described below were completed by both patients with PD and controls. The experiments were conducted in 3 sessions during 3 weeks. As we mentioned before, our thesis is part of a bigger project. The movement disorders centre will continue to recruit more patients and gather more data throughout the rest of the year. In the following sections, we describe the different scenarios and the tasks that were performed by the participants in our experiments.

4.1 Description of our experiments

In this section we describe how our experiments were conducted. The objective of our experiments is to compare how patients with Parkinson's disease carry out a series of

tasks in a virtual environment, versus how they perform the same activities in real life. Our motivation is to determine if augmented reality is suitable for rehabilitation purposes. We developed three different virtual environments. We refer to them as *scenarios*. The first scenario, called **"Watering the Plants"**, represents a living room and a kitchen. **"Supermarket"**, the second scenario, represents an aisle in a supermarket. The third scenario, **"Street Walk"**, represents a pedestrian crossing in a street. We explain these scenarios below.

4.1.1   Watering the plants scenario

This scenario represents a room filled with various combinations of flower pots. The flower pots were coloured to different colours and placed throughout this room. In this environment, subjects were asked to move toward a table where there were two rows of flower pots, one on the left and one of the right. They were then given a real watering can. The watering can, as well as the participants' hands, were segmented out to appear in the virtual world. The segmentation and overlaying gives the illusion of immersion and allow natural interaction. The patients were asked, while standing in one spot, to reach and water the furthest plant on the table in front of them, with both their right and left hands, 3 times for each hand. The patients were given the following instructions:

> "Do you see the table in front of you? There are 2 rows of plants. One row on the left and one row on the right. While standing in one spot, reach and water the furthest plant you can. Reach as far forward as possible without letting your feet raise off the ground. Hold this position until I say *STOP*. We will do this with both the right and left hand. We

will start with your right hand. Please water the plants on the right side
of the table".

It is important to mention that this procedure is performed in the virtual environment first and then in the real world. In the real world, the participant performs the same activity without the visual cues that the virtual environment provides. Figure 4.1 is a picture that shows a participant performing the "Watering the Plants" task in the virtual environment.

Figure 4.1 Watering the plants scenario: a participant performing the reach task in the virtual environment

4.1.2   Supermarket scenario

In this scenario, the participants were immersed in a three dimensional grocery store in which they could interact with augmented cereal boxes within the environment. The participants were directed to the center of the virtual room and told to face the shelve in which the cereal boxes were located. The virtual reality glasses were then put on. The participant was then directed to look at the five cereal boxes placed on top of the shelf. Then they were oriented to the room. Within the virtual room, there were five real–world numbered augmented baskets, each with a virtual sign above to indicate the number of the basket and to direct the subjects' view to the basket on the ground. The participants were then given the following instructions, which they were required to complete 3 times:

> "I am going to give you a series of 5 numbers that you will need to memorize. Each number corresponds with a basket in the room. When I say *GO*, you will need to pick up one cereal box at a time and place it in the order of baskets that I give you. The type of cereal does not matter, however the number order does. For example, if I were to give you the number sequence 1 2 3 4 5, you would walk towards the shelf, and pick up a cereal box and put it in basket 1, then you would return to the shelf pick up another cereal box and put it in basket 2, and so on. When you have completed the task, walk back to the center of the room and face the shelves"

This experiment is designed to challenge both the cognitive and motor skills of participants. The cognitive challenge is memorization. The motor challenge is requiring the participants to bend over, which is particularly difficult for patients with Parkinson's

disease. Additionally, this task in particular, helped us to observe how naturally the participants selected and manipulated the augmented objects. This procedure was performed in the virtual environment first and then in the real world. In the real world, the participants interacted with real cereal boxes and the same baskets that were visible in the virtual counterpart. The baskets were labeled in the same way as in the virtual world. Figure 4.2 shows a picture in which a participant is interacting with the augmented environment.

Figure 4.2 Supermarket Scenario: the motor and cognitive skills of participants are challenged within this task to observe gait impairment issues
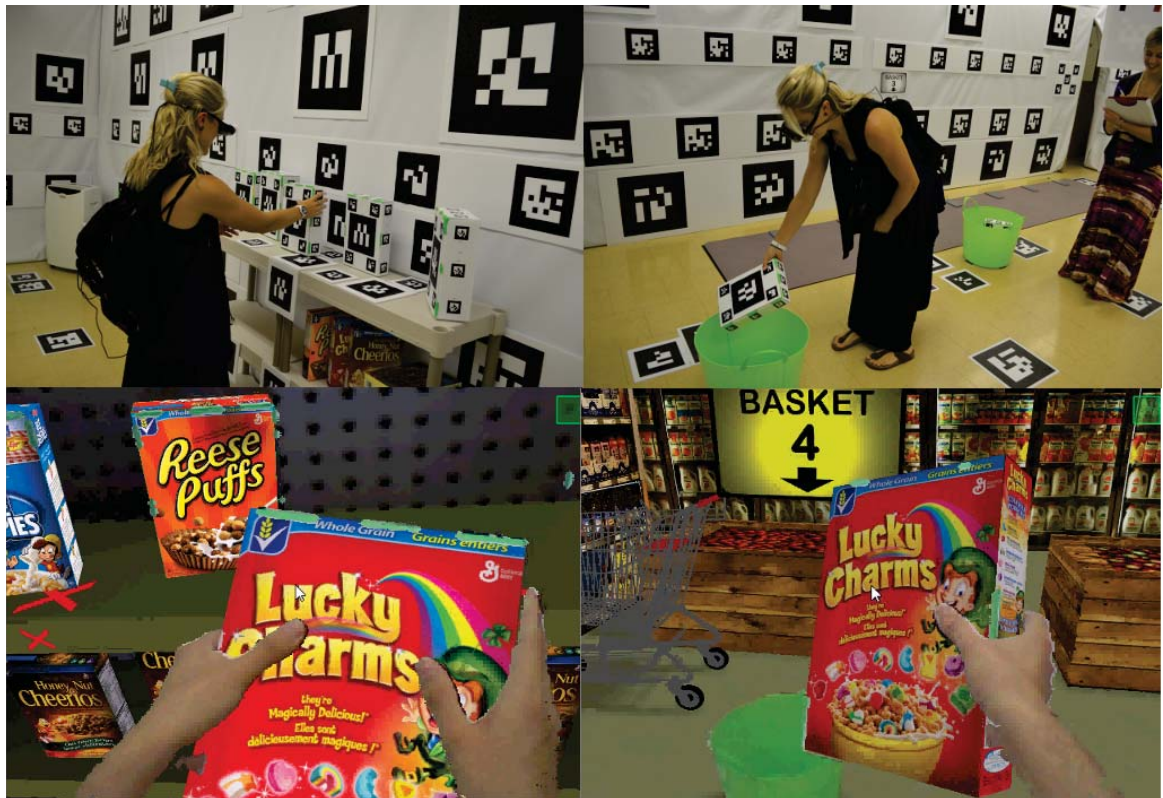
4.1.3 Street walk scenario

This scenario represents an outdoors scene, where the participants must cross the street in a crosswalk. The participants must adjust their walking speed based on instructions. The participants are asked to walk in 3 different speeds: normal speed, twice as fast with respect to their "normal" speed and half as fast with respect to their "normal" speed. Participants have to adapt their walking speed based on *internal* or *external* cues. In this context, an *internal cue* is a spoken instruction. An *external cue* is a visual element that indicates how fast the participant must walk. In this case, our external cue is a timer which displays a countdown. The participant must reach the other side of the street before the timer expires. The participants were given the following instructions:

Internal cue – normal walking speed I:

> "You are now at a crosswalk. When I say *GO*, walk across the street at a comfortable, regular pace. Keep walking until I say *LEFT*. You will then turn left and walk towards the mailbox. Please stop at the mailbox. Ready? *GO*.

Internal cue – twice as fast

> "You will now cross the street again, but this time I would like you to walk twice as fast as you just did. Keep walking until I say *LEFT*. You will then turn left and walk towards the mailbox. Please stop at the mailbox. Ready? *GO*.

Internal cue – normal walking speed II

Figure 4.3 Street walk scenario: participants have to adjust their walking speed to cross the street in the amount of time that appears in the pedestrian light

"You are now at a crosswalk. When I say *GO* walk across the street at a comfortable, regular pace. Keep walking until I say *LEFT*. You will then turn left and walk towards the mailbox. Please stop at the mailbox. Ready? *GO*.

Internal cue half as fast

"You will now cross the street again, but this time I would like you to walk half as fast as you just did. Keep walking until I say *LEFT*. You will

then turn left and walk towards the mailbox. Please stop at the mailbox.
Ready? *GO.*

We measured the time participants took to cross the street using a calibrated stopwatch. Timing began with the first step taken by participants and ended when participants reached the other side of the street. We averaged the results from **Internal cue – normal walking speed I and II** in order to obtain what we refer to as a *baseline measurement.* This baseline is our point of reference to define our external cues (i.e. the duration of the countdown). The countdown was defined as half the average baseline ($\frac{1}{2}baseline$) for the "twice as fast" trial. We defined the countdown for the "half as fast" trial as double the baseline ($2 \times baseline$). Figure 4.3 shows the outdoors scene with the external visual cue presented to participants. The instructions corresponding to the external cues are shown below.

External cue – twice as fast and half as fast

"I would like you to look at the street crossing sign in front of you. You will see an orange hand signal. When you see numbers appear on the timer below the hand, you will start walking. The timer represents the amount of time you have to cross the street. Match your walking speed to the amount of time on the timer. When you reach the end of the street, I will say *STOP.*

These same procedures were repeated in the real world. This time, in order to represent the crosswalk, we used a mat. Instead of asking the participant to walk towards the mailbox, we asked them to walk towards a red cross marked on the floor. As we mentioned before, our main variable of interest is the time to complete the different

tasks. In Chapter 5 we will present the results of our experiments. We will analyze whether there was a change in the participants' performance during the 3 weeks of trials.

Chapter 5

Experiments Results

In this chapter we present the results of the experiments that we described in Chapter 4. Our objective is to measure and compare the **time** it takes a participant to perform a series of tasks in both virtual environments and in real–life scenarios.

Our belief is that if patients take a similar amount of time to perform tasks in a virtual environment with respect to a real environment, augmented reality is not interfering with the patients' perception. Thus, the patients' experience in the augmented world can be deemed similar to the real world. This is an indication that skills learned in an augmented reality environment can be transferred to the real world.

If our beliefs were true, we could then conclude that augmented reality is adequate for the development of tools that doctors can use to assess or even rehabilitate patients.

Another way of evaluating our system is to determine how participants feel about using our system. For that reason, we asked them to complete a presence questionnaire. The objective of this questionnaire is to determine if our participants perceived our

system as realistic. Therefore, this questionnaire is valuable in assessing the suitability of our system as perceived by participants.

To evaluate participants' performance, we focus on two scenarios: *Supermarket* (see section 4.1.2) and *Street walk* (see section 4.1.3).

We do not include the results from the *Watering the plants* scenario (see section 4.1.1), because it is out of the scope of our thesis. This is because this particular scenario was not used to measure time, which is our metric of interest. Instead, it measures foot pressure using a device known as F–Scan [43]. Those measurements are being used by other researchers in order to compare participants' balance both in the virtual and real world. It is important to mention, however, that this scenario was evaluated in the presence questionnaire.

5.1 Results of the supermarket scenario experiments

As we mentioned in Chapter 4, section 4.1.2, we asked the participants to take cereal boxes and place them into baskets in an arbitrary sequence. This experiment consisted of having patients visit the hospital 3 times (1 time each week, during 3 weeks). In each visit, participants repeated the task 3 times in order to rule out measurement errors.

5.1.1 Results in the augmented reality environment

Figure 5.1 shows the evolution of the performance of the time spent by participants to complete the task during the different visits and trials using the augmented reality system. As we can observe, during the visit 1, trial 1 (`V1-T1`) while a participant took around 25 seconds to complete the task, two other participants took approximately 140 seconds. The rest of the patients' times were dispersed between 30 and 110 seconds. In that same visit, during the second trial (`V1-T2`), there was a general decrease in the time spent by participants to complete the task. For example, the participant that took 140 seconds in the previous trial (`V1-T1`), this time took around 100 seconds. This is an indication that, as users adapt to using the augmented reality system, they take less time to complete the task.

During the second visit, we can observe that time decreases further and stabilizes. It is interesting to observe that during the first visit, the range of times was between 25 and 140 seconds, while in the second visit this range was reduced to a range between 23 and 62 seconds. This tendency continues throughout the third visit.

Regarding the performance of our control participants, it is interesting to observe that one of them is very close to the mean time spent by the whole group to complete the task (refer to `VR_2C` in Figure 5.1). The second control participant (refer to `VR_1C` in Figure 5.1), took one of lowest times to complete the task.

5.1.2 Results in the real world

Figure 5.2 represents the evolution of the time it took participants to perform the Supermarket scenario task in the real world. Unlike the tendency shown in Figure

Figure 5.1 A graph showing the distribution of the times to perform tasks in the Supermarket scenario using augmented reality, during three different visits.
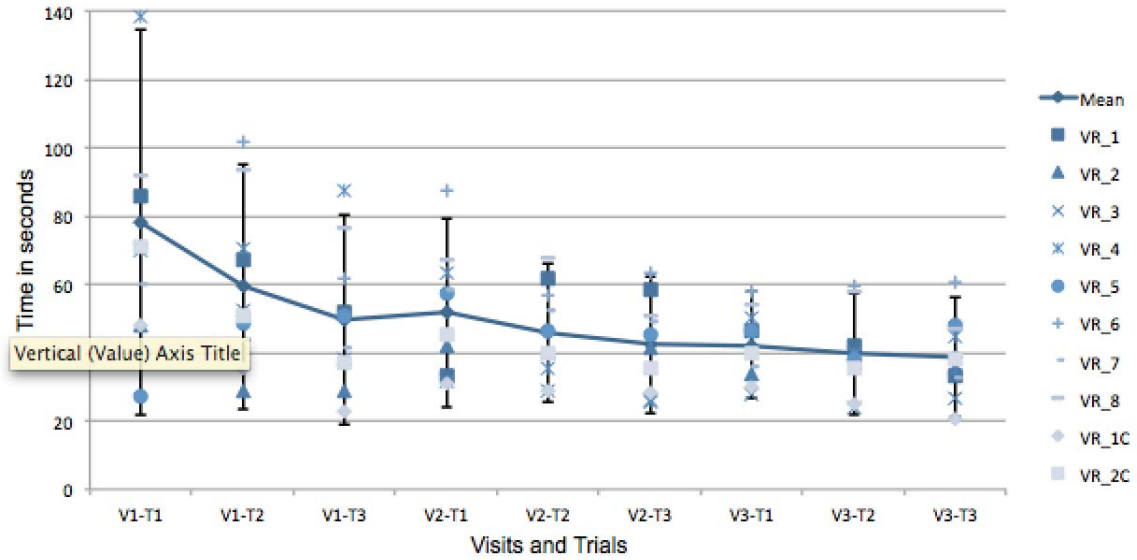
5.1 in the augmented reality experiment, the tendency for the real–world experiment is very stable and does not present a steep decrease. We believe that this is because users do not need to adapt to any system, since they are performing the task in the real world. The time range in this case is between 15 and 50 seconds.

5.1.3 Comparison of results

During the third visit in the augmented reality *Supermarket* scenario (see Figure 5.1), the performance was very close to the one observed during all 3 visits in the real–life scenario (see Figure 5.2). In spite of this similitude, in average, participants took 10 more seconds to complete the task when immersed in the augmented reality environment. This 10–second gap did not decrease throughout the rest of the visits. This

might be the result of various factors. The head mounted display was the most problematic factor in our experiments, because it considerably reduced the participants' angle of view. It also proved to be difficult to use by persons with bifocal prescription glasses.

We consider *outliers* to be participants that took an amount of time that lies outside the 1.5 standard deviation threshold. In our experiments, we found that one of our participants fell outside the upper–bound region in a normal distribution (refer to the participant VR_6 in Figures 5.1 and 5.2). We believe that the VR_6 participant performed so differently with respect to the rest of the participants, because this particular patient suffers from Parkinson's Dementia and also showed mobility issues.

Participants were able to successfully complete the task 83% of the time in the real world and 81.1% of the time in the virtual world. This indicates that augmented reality does not interfere with the participants' performance. Between all of the participants, 10 errors were made in the real world and 13 in the virtual environment. Our outlier participant, VR_6, made 35% of those errors because either he did not understand the instructions or was unable to remember the numbers provided in the instructions.

5.2   Results of the street walk scenario experiments

As we described in Chapter 4, the task in the *Street Walk* scenario consisted on asking participants to walk and adapt their walking speed according to internal and external cues. Using a baseline measurement, we asked participants to walk at two different paces: *Twice as Fast* (TF) and *Half as Fast* (HF) (see 4.1.3).

Figure 5.2 A graph showing the distribution of the times to perform tasks in the Supermarket scenario in the real world, during three different visits.
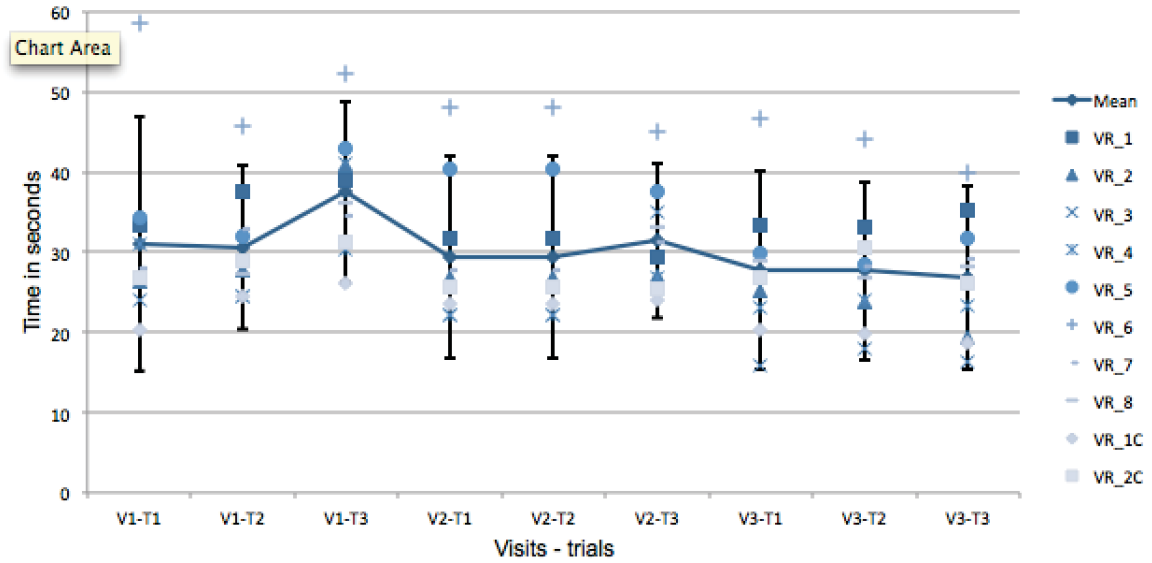
From the start, we expected that participants would adjust their walking speed better under external cues. That proved to be true. However, we were interested in measuring the difference between how well participants adapted to internal and external cues.

5.2.1 Results for the internal cue instruction

Figure 5.3 shows the distribution of the three different walking speeds during the three participant visits (one each week, during three weeks).

As we can observe, in the first visit, the interquartile range for the baseline measurement (V1-BL), was between 5.5 and 6.0 seconds. We expected that the interquartile range for (V1-TF) would be spread from 2.5 and 3.0 seconds. However, the range was

spread from 2.2 and 4.0 seconds. This means that the participants did not adapt their walking speed as expected.

In addition, we can observe that the interquartile range for the "half as fast" task in the first visit (`V1-HF`) was very similar to the baseline (`V1-BL`). This means that the majority of our participants failed to adapt their walking speed to the internal cue instruction.

During the second visit, we can observe that participants adapted their speed more efficiently with respect to the provided instructions.

During the third visit, the distribution showed that participants did not adapt their walking speed as expected. Therefore, it appears that repeating this task several times does not help participants to perform better. Nevertheless, this was expected because this task is difficult even for young, healthy people.

Figure 5.4 shows a box plot of the results that were obtained for the internal cue tasks in the real world. These results are quite similar to the results obtained in the augmented reality environment. This means that the fact that participants were not able to adapt their walking speed had nothing to do with them being in the virtual environment or the real world. Rather, this result has to do with the complexity of the task in itself.

5.2.2 Results for the external cue tasks

Figure 5.5 shows the distribution of the times that participants took to cross the street when an external cue was given. As suggest, the distributions corresponding

Figure 5.3 Box plot that shows the distribution of the three different walking speeds

Augmented Reality: Street Walk Scenario - Internal Cue comparison
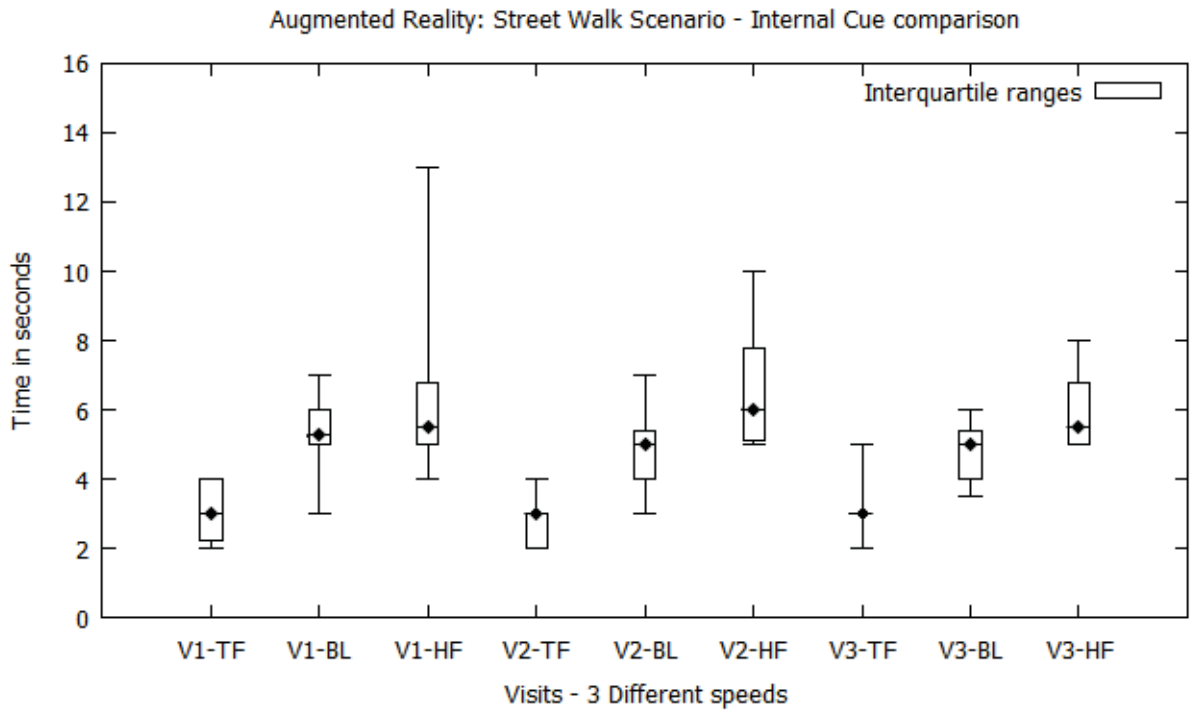
to both the "twice as fast" and "half as fast" tasks are more spread and closer to the expected values based on the baseline.

The same behavior can be observed in Figure 5.6, which shows the distributions for the real–world task when external cues were given.

5.2.3   Percentage of changes in the time to cross the street

According to the instructions that were given to the participants, they were expected to take 50% of the baseline time in the "twice as fast" task. Similarly, they were

Figure 5.4 Box plot that shows the distribution of the three different walking speeds in the real world

Real World: Street Walk Scenario - Internal Cue comparison
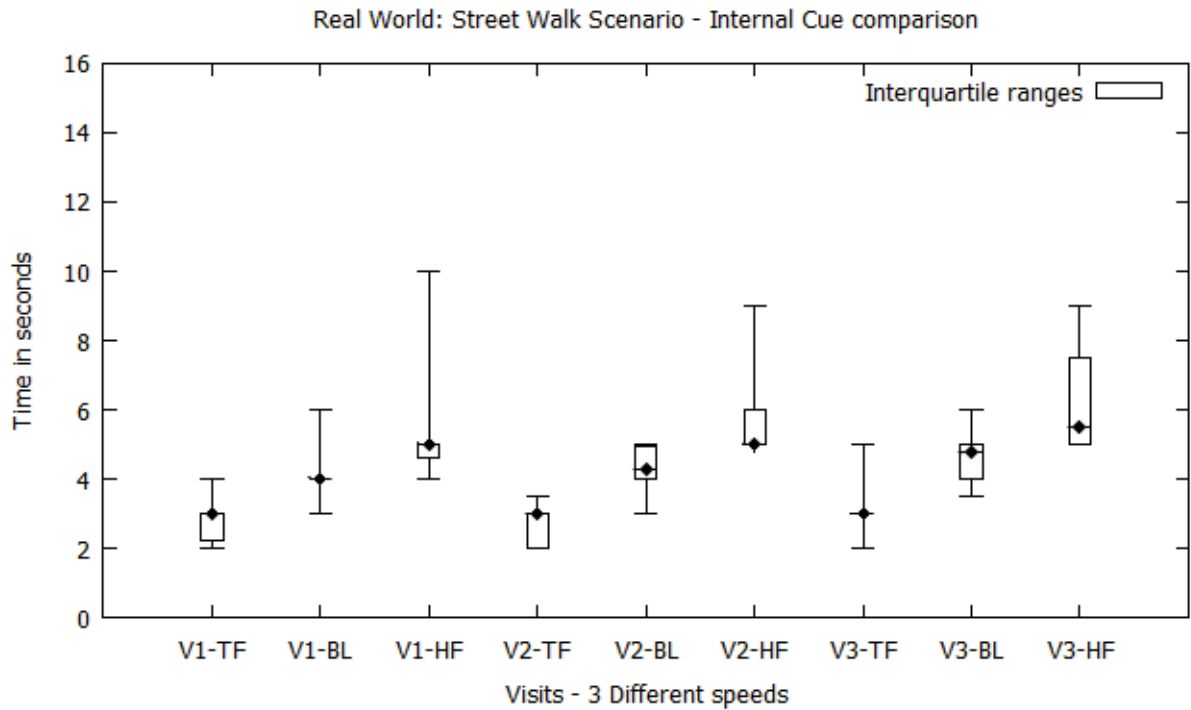
expected to take 200% of the baseline time in the "half as fast" task.

Figure 5.7 shows the average increase/decrease percentage change on the time that participants took to complete the task for the internal cue.

In the next paragraphs, we report the averaged percentage changes in walking speed for both "twice as fast" and "half as fast" tasks with internal cues.

Twice as fast task – internal cue. The results that we obtained show that participants increased their speed and took 60% of the baseline time when they were

Figure 5.5 Box plot that shows the distribution of the three different walking speeds in the augmented reality for the external cue tasks
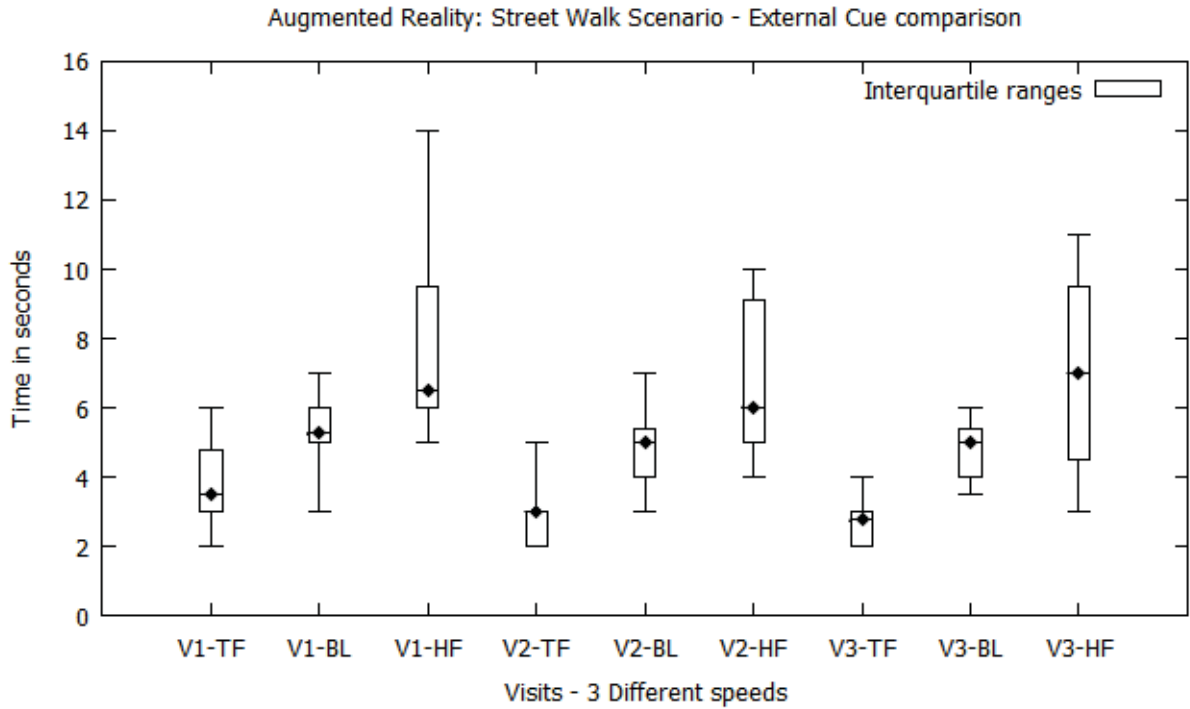
given the internal cue in the augmented environment. Conversely, the participants increased their speed and took 65% of the baseline time when they performed the task in the real world. In this specific task, the participants adapted their walking speed better in the augmented environment.

Half as fast task – internal cue. The results that we obtained show that participants decreased their speed and took 131% of the original baseline time in the augmented environment. In contrast, the participants decreased their speed and took 136% of the baseline time when they performed the task in the real world. These results are very similar between them and therefore, we can observe that augmented

Figure 5.6 Box plot that shows the distribution of the three different walking speeds in the real world for the external cue tasks
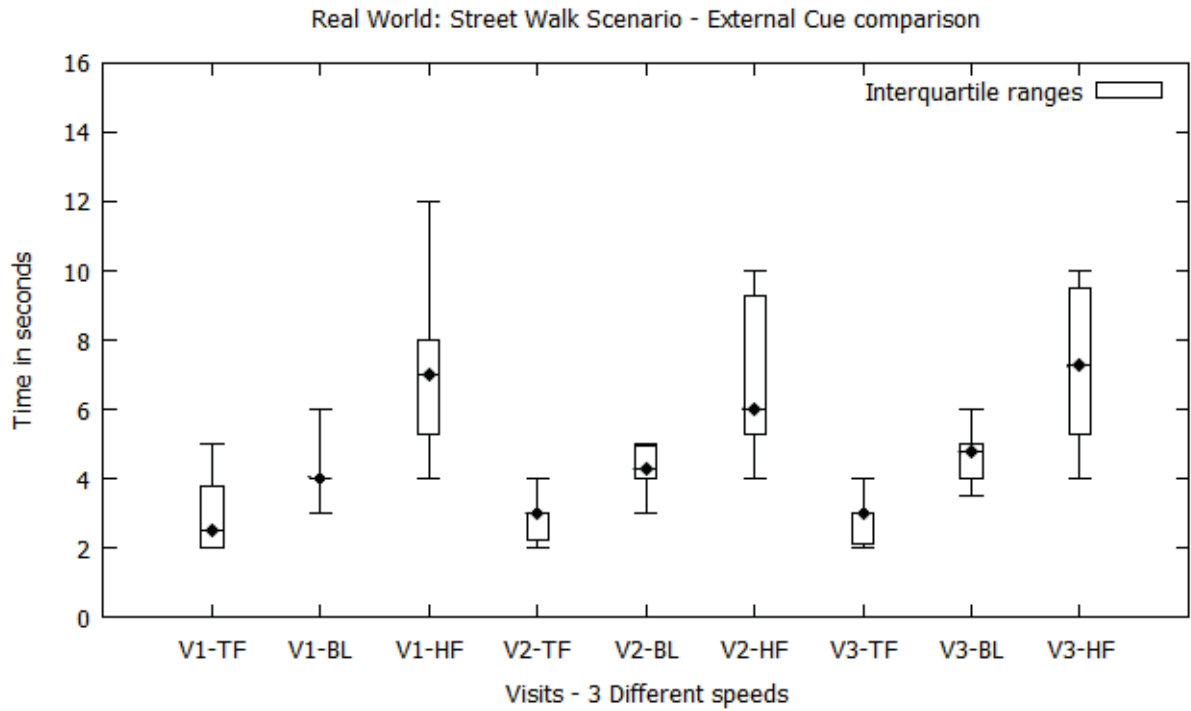
reality did not affect the results.

Figure 5.8 shows the average increase/decrease percentage change on the time that participants took to complete the task for the external cue.

In the next paragraphs, we present the averaged percentage changes in walking speed for both "twice as fast" and "half as fast" tasks with external cues.

Twice as fast task – external cue. Our measurements show that participants increased their speed and took 62% of the baseline time when the external cue was shown in the augmented environment. Conversely, the participants increased their

Figure 5.7 Table: average increase/decrease percentage change on the time that participants took to complete the task with internal cues

	INTERNAL CUE			
	TWICE AS FAST TASK (Average per participant throughout 3 visits)		**HALF AS FAST TASK** (Average per participant throughout 3 visits)	
Participants	Augmented Reality	Real World	Augmented Reality	Real World
VR_1	59%	73%	131%	136%
VR_2	63%	57%	124%	140%
VR_3	47%	61%	117%	153%
VR_4	59%	62%	113%	132%
VR_5	67%	75%	158%	147%
VR_6	60%	66%	89%	101%
VR_7	67%	75%	108%	121%
VR_8	67%	62%	143%	169%
VR_1C	61%	64%	178%	139%
VR_2C	48%	58%	152%	125%
Avg. for all participants	**60%**	**65%**	**131%**	**136%**

speed and took 66% of the baseline time when they performed the task in the real world. The results for the external cue, twice as fast task are consistent with the internal cue results. The participants adjusted their speed better when the task was performed in the augmented environment.

Half as fast task – external cue. The results show that participants decreased their speed and took 148% of the original baseline time in the augmented environment with external cues. In contrast, the participants decreased their speed and took 160% of the baseline time when they performed the task in the real world. These results show that participants adapted their walking speed better when they performed the task in the real world.

Figure 5.8 Table: average increase/decrease percentage change on the time that participants took to complete the tasks with the external cue

	EXTERNAL CUE			
	TWICE AS FAST TASK (Average per participant throughout 3 visits)		HALF AS FAST TASK (Average per participant throughout 3visits)	
Participants	Augmented Reality	Real World	Augmented Reality	Real World
VR_1	53%	53%	137%	158%
VR_2	57%	57%	191%	200%
VR_3	53%	72%	123%	131%
VR_4	64%	78%	130%	138%
VR_5	50%	72%	125%	119%
VR_6	77%	81%	88%	96%
VR_7	63%	75%	92%	150%
VR_8	76%	63%	181%	200%
VR_1C	83%	56%	200%	196%
VR_2C	48%	53%	217%	217%
Avg. for all participants	62%	66%	148%	160%

5.2.4   Comparison between the real and the virtual worlds

In this scenario, we found that there is not a significant difference between the performance of participants in the real and virtual worlds. In some tasks, participants performed better in the virtual environment while in other tasks, there is a slight increase in performance in the real world. The maximum difference in performance between the virtual and real worlds is 5%. We do not consider that difference to be important. Therefore, we found that for this specific activity augmented reality is not interfering with the task and that the participants' experience was similar in both cases.

5.3   Presence questionnaire evaluation

The effectiveness of a virtual environment has been linked to the sense of presence reported by the user. *Presence* can be defined as a normal awareness phenomenon that requires attention and is based in the interaction between sensory stimulation, environmental factors and internal tendencies to become involved [66].

To evaluate if our augmented reality system provided an adequate level of presence enough to immerse patients in the different scenarios, we asked our participants to complete a presence questionnaire after they finished the tasks in every visit. The presence questionnaire we employed is based on the work by Witmer and Singer [66].

The presence questionnaire consists of 34 questions with a 7–point Likert scale, which evaluates different factors that affect the involvement of participants in a virtual environment and thus the level of immersion. These factors can be classified in four categories: control factors, distraction factors, sensory factors and realism factors. *Control factors* refer to the degree of control a user can have when interacting with a virtual environment. In addition, this factor evaluates if a user gets satisfactory feedback to an action. *Sensory factors* refer to how many senses are involved when using a system. *Distraction factors* evaluate whether the hardware interferes with the degree of focus that users achieve.

Realism factors measure how well the virtual environment is built to simulate real world places.

Participants evaluated the system by selecting a box of the scale in accordance with the question content and the descriptive labels. We assigned a value from 1 to 7 depending on the box they selected, being 7 the most positive answer regarding our

system and 1 being the most negative. Figure 5.9 shows an example with the type of questions contained in the presence questionnaire.

Figure 5.9 Example of the questions included in the presence questionnaire

Presence Questionnaire

Please place an "X" in the appropriate box of the scale in accordance with the question content and descriptive labels.

1. How much were you able to control events?

| NOT ABLE | MODERATELY ABLE | VERY ABLE |

2. How responsive was the environment to actions that you initiated (or performed)?

| NOT RESPONSIVE | MODERATELY RESPONSIVE | VERY RESPONSIVE |

3. How natural did you interactions with the environment seem?

| NOT NATURAL | MODERATELY NATURAL | VERY NATURAL |

Figure 5.10 shows a table that shows the average rating that users gave to the system. This average rating was obtained by averaging the answers to the 34 questions for each participant.

In general, we observe that participants rated their experience as moderately real to very good and excellent. From this, we can conclude that participants had an overall favorable perception of the system.

Figure 5.10 Presence questionnaire results

Average from 34 questions				
Participant	Visit 1	Visit 2	Visit 3	Average of 3 visits
VR_X	3.106	3.000	3.212	3.106
VR_1	5.636	5.212	5.545	5.465
VR_2	6.182	6.061	6.333	6.192
VR_3	4.909	4.576	4.970	4.818
VR_4	6.182	6.273	6.091	6.182
VR_5	5.364	4.788	4.424	4.859
VR_6	4.061	4.727	4.545	4.444
VR_7	5.970	5.848	5.545	5.788
VR_8	5.212	5.818	5.667	5.566
VR_1C	6.091	5.909	6.000	6.000
VR_2C	5.515	5.091	5.364	5.323

5.3.1 Relationship between sense of presence and participants' performance

We are also interested in determining if there is a relationship between how the participants evaluated the system and how well they performed in our experiments. For example, the participant (VR_X) that gave the lowest grade to our system (3.1) was severely affected by dementia and failed to complete the tasks. This participant did try to complete the tasks through our system. However, the HMD made it difficult because it limited the participants' field of view. In addition, this participant suffered from a slouched posture. This, along with the HMD's reduced field of vision made it very difficult for this participant to see the virtual world adequately. For these reasons, we do not have task–related data for this participant. However, we did ask him to complete the presence questionnaire.

Another example of a participant that gave a low score to the system was the partic-

ipant `VR_6`, who did not perform well in the tasks (see Figures 5.1 and 5.2).

The two participants who gave the highest scores were `VR_2` and `VR_4`. As we can observe in Figure 5.1, both `VR_2` and `VR_4` had a good performance. In addition, participant `VR_4` showed the best performance evolution throughout time.

5.4 Discussion

Our experiments helped us understand the benefits and limitations of immersive augmented reality. Overall, participants had a positive opinion regarding our system, as reflected by the presence questionnaire. Although there was a difference between the time participants took to complete tasks in the *Supermarket* augmented reality environment against the real world, participants were able to successfully complete the tasks *in both cases*. Regarding the *Street Walk* scenario, our results show that the performance of participants in this task was very similar both in the augmented reality environment and real world. Thus, we can conclude that augmented reality was not a factor in the performance of our participants.

We were able to successfully develop an augmented reality system that allows people to freely navigate virtual environments. Moreover, it allows natural interaction with both real and augmented objects. Therefore, this system can be used not only as a support tool in rehabilitation programs, but in other areas such as industrial training.

The main limitation of our system is the head mounted display. For example, we found out that the head mounted display limited the participants' field of view, which affected their perception of the virtual environment. Basically, the HMD eliminates peripheral vision. The head mounted display we used for our experiments, provides

only 32 degrees of vertical field of view. This is a huge limitation compared to the normal human eye vertical field of view of 120 degrees [4]. Because of the visual limitation of the current HMD, we observed that augmented reality systems are not suitable for people that suffer from slouched posture. This is because they cannot see important aspects of the virtual world through the head mounted display. The HMD limited field of view combined with some participants' slouched posture caused them to only be able to see the floor of the virtual environment. We are studying the possibility of using a different head mounted display with a wider field of view for future studies.

Chapter 6

Conclusions

6.1   Conclusion

The main objective of this thesis was to develop and evaluate an augmented reality system. We successfully developed a flexible augmented reality system that could be used by doctors as a tool to assess their patients. We consider that one of the most important contributions of our system is that it provides users with the ability to naturally interact with objects without the need for external devices. For example, instead of interacting with the system through a mouse or glove, users are able to grab objects with their own hands. This was made possible by our implementation of a skin classification algorithm that allows our system display the users' hands on top of the virtual environment.

One of the most important features of our system is that it provides free navigation. That means that users can walk and move freely within the virtual environment, as

they would do in a real–life situation. Regarding navigation, the only limitation of our system is determined by the physical space where the system is deployed. Free navigation was implemented by using vision–based tracking. This feature allows users to feel as if they were *actually present* in the virtual environment by providing a first–person view. Another objective of this thesis was to perform experiments in order to determine if augmented reality can be used in future rehabilitation applications.

Our experiments consisted in comparing how users performed a series of tasks in a virtual environment versus how users performed the same tasks in the real–world. To perform this comparison, we measured the time it took users to complete a set of predefined tasks. Our belief is that if a user takes more time to perform a task in a virtual environment with respect to the real world, then that could indicate that the user might be having difficulty perceiving the virtual environment as real.

Our results show that the time it took participants to complete tasks in the augmented world is very similar to the time it takes to complete the same tasks in the real world. From this, we can conclude that augmented reality provides a realistic environment where users can perform tasks in a similar way as they would do in real life. Also, we found that there is a relation between the sense of presence that participants experimented, and how well they performed in tasks in the augmented environment. This means that if people perceive the virtual environment as being "natural", there are more possibilities to obtain attention and learning that can be transferred to real world activities.

During our experiments, we were able to identify two limitations in our system. One limitation is that the head mounted display we used is not designed to be worn over prescription glasses. Another limitation was that the limited angle of view that the head mounted display provides, limited users' peripheral vision. This caused them to

be unable to see their own hands at times.

We believe that augmented reality can be successfully applied not only in medicine applications, but also in training applications.

6.2 Future work

In order to setup the physical space needed for the system, it is necessary to manually configure all of the fiducial markers to be used. It would be desirable to implement a feature where users could use fewer markers and complement tracking by using natural features to reduce the time needed to setup the system.

Regarding hand and objects of interest segmentation, it is necessary to strengthen our classification algorithms. As future work, we propose to mix our algorithms with depth perception and a object segmentation to allow multiple levels of occlusion between the real world and the virtual environment.

In order to confirm if augmented reality can be used for rehabilitation programs, further experimental work is needed to gather further feedback that help us to improve and refine the system.

Bibliography

[1] Open source computer vision library. `http://sourceforge.net/projects/opencvlibrary/`. Accessed: July 2012.

[2] J.F. Aita. Why patients with parkinson's disease fall. *JAMA: the journal of the American Medical Association*, 247(4):515–516, 1982.

[3] G. Albani, R. Pignatti, L. Bertella, L. Priano, C. Semenza, E. Molinari, G. Riva, and A. Mauro. Common daily activities in the virtual environment: a preliminary study in parkinsonian patients. *Neurological Sciences*, 23:49–50, 2002.

[4] K. Arthur. Effects of field of view on task performance with head-mounted displays. In *Conference companion on Human factors in computing systems: common ground*, pages 29–30. ACM, 1996.

[5] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre. Recent advances in augmented reality. *Computer Graphics and Applications, IEEE*, 21(6):34–47, 2001.

[6] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre. Recent advances in augmented reality. *Computer Graphics and Applications, IEEE*, 21(6):34–47, 2001.

[7] G. Baratoff and S. Blanksteen. Tracking devices. *Encyclopedia of Virtual Environments*, 1993.

[8] W. Barfield, D. Zeltzer, T. Sheridan, and M. Slater. Presence and performance within virtual environments. *Virtual environments and advanced interface design*, pages 473–513, 1995.

[9] MJA Berry and G. Linoff. Data mining techniques: for marketing, sales, and customer support 1997. *John Willey & Sons*, 1997.

[10] M. Billinghurst, H. Kato, and I. Poupyrev. Collaboration with tangible augmented reality interfaces. In *HCI International*, volume 1, pages 5–10, 2001.

[11] E. Bisson, B. Contant, H. Sveistrup, and Y. Lajoie. Functional balance and dual-task reaction times in older adults are improved by virtual reality and biofeedback training. *Cyberpsychology & Behavior*, 10(1):16–23, 2007.

[12] J. Brand and J.S. Mason. A comparative assessment of three approaches to pixel-level human skin-detection. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, volume 1, pages 1056 –1059 vol.1, 2000.

[13] V. Buchmann, S. Violich, M. Billinghurst, and A. Cockburn. Fingartips: gesture based direct manipulation in augmented reality. In *Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and South East Asia*, pages 212–221. ACM, 2004.

[14] JJ Chen. Parkinson's disease: health-related quality of life, economic cost, and implications of early treatment. *The American journal of managed care*, 16:S87, 2010.

[15] H. Creagh. Cave automatic virtual environment. In *Electrical Insulation Conference and Electrical Manufacturing Coil Winding Technology Conference, 2003. Proceedings*, pages 499 – 504, sept. 2003.

[16] T. Darrell, G. Gordon, M. Harville, and J. Woodfill. Integrated person tracking using stereo, color, and pattern detection. In *Computer Vision and Pattern Recognition, 1998. Proceedings. 1998 IEEE Computer Society Conference on*, pages 601 –608, jun 1998.

[17] S. Davidsdottir, R. Wagenaar, D. Young, and A. Cronin-Golomb. Impact of optic flow perception and egocentric coordinates on veering in parkinson's disease. *Brain*, 131(11):2882–2893, 2008.

[18] N. Durlach and M. Slater. Presence in shared virtual environments and virtual togetherness. *Presence: Teleoperators & Virtual Environments*, 9(2):214–217, 2000.

[19] JW Eaton, D. Bateman, and S. Hauberg. Gnu octave manual version 3 (network theory ltd., united kingdom, 2008).

[20] A. Elgammal, C. Muang, and D. Hu. Skin detection-a short tutorial. *Encyclopedia of Biometrics RU Department of Computer Science. Piscataway, NJ*, 8902:1–10, 2009.

[21] B. Furht. *Handbook of augmented reality*. Springer-Verlag New York Inc, 2011.

[22] Francesca Gasparini and Raimondo Schettini. Skin segmentation using multiple thresholding. *Internet Imaging from International Society of Optics and Photonics, proceedings of*, I(7):862 –877, january 2006.

[23] E.B. Goldstein. *Sensation and perception*. Wadsworth Pub Co, 2009.

[24] C. Hendrix and W. Barfield. The sense of presence within auditory virtual environments. *Presence Teleoperators Virtual Environments*, 5(3):290–301, 1996.

[25] J.H. Hollman, R.H. Brey, R.A. Robb, T.J. Bang, and K.R. Kaufman. Spatiotemporal gait deviations in a virtual reality environment. *Gait & posture*, 23(4):441–444, 2006.

[26] Y. Huang, Y. Liu, and Y. Wang. Ar-view: An augmented reality device for digital reconstruction of yuangmingyuan. In *Mixed and Augmented Reality-Arts, Media and Humanities, 2009. ISMAR-AMH 2009. IEEE International Symposium on*, pages 3–7. Ieee, 2009.

[27] D.M. Huse, K. Schulman, L. Orsini, J. Castelli-Haley, S. Kennedy, and G. Lenhart. Burden of illness in parkinson's disease. *Movement disorders*, 20(11):1449–1454, 2005.

[28] H. Ishii. Ullmer., b.(1997). tangible bits: Towards seamless interfaces between people, bits and atoms. In *Proceedings of the ACM conference on Human Factors in Computing Systems*, pages 234–241.

[29] H. Iwata. Walking about virtual environments on an infinite floor. In *Virtual Reality, 1999. Proceedings., IEEE*, pages 286–293. IEEE, 1999.

[30] J. Jankovic. Parkinson's disease: clinical features and diagnosis. *Journal of Neurology, Neurosurgery & Psychiatry*, 79(4):368, 2008.

[31] M.J. Jones and J.M. Rehg. Statistical color models with application to skin detection. In *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, volume 1, pages 2 vol. (xxiii+637+663), 1999.

[32] T.A. Kaminsky, B.J. Dudgeon, F.F. Billingsley, P.H. Mitchell, and S.J. Weghorst. Virtual cues and functional mobility of people with parkinson's disease: A single-subject pilot study. *Journal of Rehabilitation Research and Development*, 44(3):437, 2007.

[33] Garret Kapp. The application of virtual reality for parkinson's disease rehabilitation. Undergraduate Thesis, 2010.

[34] H. Kato and M. Billinghurst. Marker tracking and hmd calibration for a video-based augmented reality conferencing system. In *Augmented Reality, 1999.(IWAR'99) Proceedings. 2nd IEEE and ACM International Workshop on*, pages 85–94. IEEE, 1999.

[35] H. Kato, M. Billinghurst, I. Poupyrev, K. Imamoto, and K. Tachibana. Virtual object manipulation on a table-top ar environment. In *Augmented Reality, 2000.(ISAR 2000). Proceedings. IEEE and ACM International Symposium on*, pages 111–119. Ieee, 2000.

[36] H. Kato, K. Tachibana, M. Tanabe, T. Nakajima, and Y. Fukuda. Magiccup: A tangible interface for virtual objects manipulation in table-top augmented reality. In *Augmented Reality Toolkit Workshop, 2003. IEEE International*, pages 75–76. IEEE, 2003.

[37] Georg Klein. *Visual Tracking for Augmented Reality*. Monography. Department of Engineering.

[38] R. Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *International joint Conference on artificial intelligence*, volume 14, pages 1137–1145. LAWRENCE ERLBAUM ASSOCIATES LTD, 1995.

[39] G. Kwakkel, CJT De Goede, and EEH Van Wegen. Impact of physical therapy for parkinson's disease: a critical review of the literature. *Parkinsonism & related disorders*, 13:S478–S487, 2007.

[40] B. Lee and J. Chun. Interactive manipulation of augmented objects in markerless ar using vision-based hand interaction. In *Information Technology: New Generations (ITNG), 2010 Seventh International Conference on*, pages 398–403. IEEE, 2010.

[41] G.A. Lee, C. Nelles, M. Billinghurst, and G.J. Kim. Immersive authoring of tangible augmented reality applications. In *Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality*, pages 172–181. IEEE Computer Society, 2004.

[42] J.H. Lee, J. Ku, W. Cho, W.Y. Hahn, I.Y. Kim, S.M. Lee, Y. Kang, D.Y. Kim, T. Yu, B.K. Wiederhold, et al. A virtual reality system for the assessment and rehabilitation of the activities of daily living. *CyberPsychology & Behavior*, 6(4):383–388, 2003.

[43] Z.P. Luo, L.J. Berglund, K.N. An, et al. Validation of f-scan pressure sensor system: a technical note. *Journal of rehabilitation research and development*, 35:186–186, 1998.

[44] A.S. Merians, D. Jack, R. Boian, M. Tremaine, G.C. Burdea, S.V. Adamovich, M. Recce, and H. Poizner. Virtual reality–augmented rehabilitation for patients following stroke. *Physical therapy*, 82(9):898–915, 2002.

[45] J. Messier, S. Adamovich, D. Jack, W. Hening, J. Sage, and H. Poizner. Visuomotor learning in immersive 3d virtual reality in parkinsons disease and in aging. *Experimental brain research*, 179(3):457–474, 2007.

[46] A. Mirelman, I. Maidan, T. Herman, J.E. Deutsch, N. Giladi, and J.M. Hausdorff. Virtual reality for gait training: can it induce motor learning to enhance complex walking and reduce fall risk in patients with parkinson's disease? *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 66(2):234, 2011.

[47] T.M. Mitchell. Machine learning. wcb. *Mac Graw Hill*, page 368, 1997.

[48] M. Moehring and B. Froehlich. Enabling functional validation of virtual cars through natural interaction metaphors. In *Virtual Reality Conference (VR), 2010 IEEE*, pages 27–34. IEEE, 2010.

[49] J. Nilsson, A.C.E. Odblom, J. Fredriksson, A. Zafar, and F. Ahmed. Performance evaluation method for mobile computer vision systems using augmented reality. In *Virtual Reality Conference (VR), 2010 IEEE*, pages 19 –22, march 2010.

[50] J.A. O'Brien, A. Ward, SL Michels, S. Tzivelekis, and NJ Brandt. Economic burden associated with parkinson disease. *Drug Benefit Trends*, 21(6):179–190, 2009.

[51] C.B. Owen, Fan Xiao, and P. Middlin. What is the best fiducial? In *Augmented Reality Toolkit, The First IEEE International Workshop*, page 8 pp., 2002.

[52] Son Lam Phung, D. Chai, and A. Bouzerdoum. Adaptive skin segmentation in color images. In *Multimedia and Expo, 2003. ICME '03. Proceedings. 2003 International Conference on*, volume 3, pages III – 173–6 vol.3, july 2003.

[53] Warren S. Sarle, 2002. Artificial Neural Neworks FAQ.

[54] L.F. Schettino, S.V. Adamovich, W. Hening, E. Tunik, J. Sage, and H. Poizner. Hand preshaping in parkinsons disease: effects of visual feedback and medication state. *Experimental brain research*, 168(1):186–202, 2006.

[55] L. Sigal, S. Sclaroff, and V. Athitsos. Skin color-based video segmentation under time-varying illumination. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(7):862 –877, july 2004.

[56] M. Slater and S. Wilbur. A framework for immersive virtual environments (five): Speculations on the role of presence in virtual environments. *Presence Teleoperators and Virtual Environments*, 6(6):603–616, 1997.

[57] JL Souman, P.R. Giordano, M. Schwaiger, I. Frissen, T. Thümmel, H. Ulbrich, A.D. Luca, HH Bülthoff, and M.O. Ernst. Cyberwalk: Enabling unconstrained omnidirectional walking through virtual environments. *ACM Transactions on Applied Perception (TAP)*, 8(4):25, 2011.

[58] M. Storring, H.J. Andersen, and E. Granum. Estimation of the illuminant colour from human skin colour. In *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*, pages 64 –69, 2000.

[59] J.N. Templeman, P.S. Denbrook, and L.E. Sibert. Virtual locomotion: Walking in place through virtual environments. *Presence*, 8(6):598–617, 1999.

[60] J.-C. Terrillon, M.N. Shirazi, H. Fukamachi, and S. Akamatsu. Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images. In *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*, pages 54 –61, 2000.

[61] B. Thomas, B. Close, J. Donoghue, J. Squires, P. De Bondi, M. Morris, and W. Piekarski. Arquake: An outdoor/indoor augmented reality first person application. In *Wearable Computers, 2000. The Fourth International Symposium on*, pages 139–146. IEEE, 2000.

[62] B.H. Thomas and W. Piekarski. Glove based user interaction techniques for augmented reality in an outdoor environment. *Virtual Reality*, 6(3):167–180, 2002.

[63] Vladimir Vezhnevets, Vassili Sazonov, and Alla Andreeva. A survey on pixel-based skin color detection techniques. In *IN PROC. GRAPHICON-2003*, pages 85–92, 2003.

[64] Vedran Vujcic. Virtual reality for the use of medical research. Undergraduate Thesis, 2010.

[65] B. Williams, S. Bailey, G. Narasimham, M. Li, and B. Bodenheimer. Evaluation of walking in place on a wii balance board to explore a virtual environment. *ACM Transactions on Applied Perception (TAP)*, 8(3):19, 2011.

[66] B.G. Witmer and M.J. Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence*, 7(3):225–240, 1998.

[67] Matthias M. Wloka and Brian G. Anderson. Resolving occlusion in augmented reality. In *Proceedings of the 1995 symposium on Interactive 3D graphics*, I3D '95, pages 5–12, New York, NY, USA, 1995. ACM.

[68] D.E. Young, R.C. Wagenaar, C.C. Lin, Y.H. Chou, S. Davidsdottir, E. Saltzman, and A. Cronin-Golomb. Visuospatial perception and navigation in parkinsons disease. *Vision research*, 50(23):2495–2504, 2010.

[69] L. Zhang, B.C. Abreu, G.S. Seale, B. Masel, C.H. Christiansen, and K.J. Ottenbacher. A virtual reality environment for evaluation of a daily living skill in brain injury rehabilitation: reliability and validity1. *Archives of physical medicine and rehabilitation*, 84(8):1118–1124, 2003.

[70] F. Zhou, H.B.L. Duh, and M. Billinghurst. Trends in augmented reality tracking, interaction and display: A review of ten years of ismar. In *Mixed and Augmented Reality, 2008. ISMAR 2008. 7th IEEE/ACM International Symposium on*, pages 193–202. Ieee, 2008.

[71] F. Zhou, H.B.L. Duh, and M. Billinghurst. Trends in augmented reality tracking, interaction and display: A review of ten years of ismar. In *Mixed and Augmented Reality, 2008. ISMAR 2008. 7th IEEE/ACM International Symposium on*, pages 193–202. IEEE, 2008.

Vita

Name	Andrés Ayala García
Place of Birth	Mexico City, Mexico
Year of Birth	1981
Post-secondary Education and Degrees	The University of Western Ontario London, Ontario, Canada 2011–2012 M. Sc. in Computer Science Zacatepec Institute of Technology Morelos, Mexico 1999–2004 B.Eng. in Computer Software
Honours and Awards:	CONACYT International Graduate Scholarship, 2011–2012
Related work experience:	Research Assistant London Health Sciences Centre Ontario, Canada 2012 Software Developer/Project Manager The Electric Research Institute Cuernavaca, Mexico 2005–2010 Software Developer/Research Assistant The Institute of Photonics and Electronics, Academy of Sciences CR. Prague, Czech Republic 2004