Out-of-Reach Interactions in VR

Eduardo Augusto de Libório Cordeiro

Thesis to obtain the Master of Science Degree in

Computer Engineering

Supervisor(s): Prof. Alfredo Manuel dos Santos Ferreira Júnior
Eng. Daniel Filipe Martins Tavares Mendes

Examination Committee

Chairperson: Prof. Miguel Nuno Dias Alves Pupo Correia
Supervisor: Prof. Alfredo Manuel dos Santos Ferreira Júnior
Member of the Committee: Prof. Maria Teresa Caeiro Chambel

November 2016

Dedicated to my parents.

Acknowledgments

I have to deeply thank the support, encouragement and incentive that was given to me during this dissertation. First of all, to Instituto Superior Técnico for the opportunity to work in such an engaging topic, and also to the VIMMI group of INESC-ID Lisboa for being so welcoming. To Prof. Doutor Alfredo Ferreira for all the guidance and advice during this challenge.

I would also like to express gratitude to Daniel Mendes, the co-advisor for this thesis. I have an immense respect for all the hard work he has done throughout this period. Without his suggestions my thesis would not be as valuable, and also for his never ending fortitude during the most complicated chapters of this work.

To my friend Filipe Relvas, for the many discussions, problem solving conversations and for all the help given. Also a thank you to Daniel Medeiros for all the guidance and patience provided, especially during the most demanding periods. A big thanks to Mauricio Sousa for all the assistance, and also for providing many needed moments of entertainment.

I would also like to extend my thanks to everyone that gladly participated in the user evaluation sessions.

I would also like to mention the financial support from the Portuguese Foundation for Science and Technology (FCT), through projects TECTON3D (ref. PTDC/EEI-SII/3154/2012) and IT-MEDEX (ref. PTDC/EEISII/6038/2014).

And a special thanks to my parents, for the constant support and encouragement, and without whom I would have not achieved this goal. I would also like to thank them for being so understanding and helping me through all my life.

Resumo

O acto de apontar para um objecto é fundamental no mundo real. Quando interagimos com objetos virtuais, esta acção é considerada como a seleção de um objecto. Existem várias abordagens, usadas desde *tabletops* até aos ambientes virtuais. No entanto, quando os objectos estão muito distantes do utilizador as abordagens actuais não são capazes de proporcionar níveis de velocidade e precisão adequados. Abordagens para ultrapassar estas limitações destas técnicas frequentemente favorecem o uso de volumes de selecção, em oposição das que fazem uso de *ray-casting* ou *arm-extension*. Mesmo com o uso de volumes de selecção, a selecção de objectos distantes ainda é um desafio, pois estes volumes estão propensos a selecções incorrectas.

O nosso foco vai ser sobre este desafio. Foi feita pesquisa sobre técnicas de selecção, e que foram depois comparadas usando a nossa taxonomia. Para superar os desafios abertos das técnicas de selecção estudadas, desenvolvemos a técnica PRECIOUS, uma técnica de seleção de objectos fora de alcance. Esta técnica combina refinamento progressivo com *cone casting* e é capaz de seleccionar objectos a distancias variadas. Usando uma técnica de viagem, escolhida com base numa avaliação feita entre três técnicas, o utilizador é movido para mais perto do objecto que pretende seleccionar.

Esta técnica foi comparada com outras abordagens para a selecção de objectos distantes. Os resultados indicam que mesmo não proporcionando os tempos de selecção mais rápidos, a ausência significativa de selecções errada e a uniformidade nos tempos de selecção em vários cenários, fazem dela uma abordagem apropriada para a selecção de objectos fora de alcance.

Palavras-chave: Selecção de Objectos Fora de Alcance, Avaliação de Técnicas de Viagem, Concepcão Centrada no Utilizador, Realidade Virtual

Abstract

The act of pointing to an object is fundamental in our everyday life. When interacting with virtual objects, this action is called object selection. There are several approaches to object selection, which are used in tabletops all the way to virtual environments. However, when objects are very distant from the user, the current approaches are not able to provide adequate levels of speed and accuracy. Approaches to surpass the limitations of these techniques often favour the use of selection volumes to ray-casting and arm-extension. Even with the use of selection volumes, distant object selection still poses a challenge, as these volumes are prone to unwanted selections.

We will focus on this challenge.

We surveyed several selection techniques which were then compared based on our taxonomy. To overcome the open challenges of the selection techniques studied, we developed PRECIOUS, a selection technique for out-of-reach selection. It combines progressive refinement with cone casting and is capable of selecting objects at various distances. Using a travel technique, chosen based of an evaluation conducted between three techniques, we bring the user closer to the object he intends to select.

PRECIOUS was then compared with other approaches on distant object selection. The results indicate that even though it does not provide the faster completion times, the significant absence of incorrect selections and the consistent completion times across different scenarios make it an appropriate approach to out-of-reach selection.

Keywords: Out-of-reach Selection, Travel Technique Evaluation, User-Centered Design, Virtual Reality

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Immersive Virtual Environments (IVEs) have been the object of considerable interest in the latest years. The development of new hardware has allowed for the development of techniques and applications that can make use of these environments.

In recent years there has been an accelerated advance in virtual reality technology, mainly with the appearance of new low-cost stereoscopic hardware such as the Oculus Rift [1] and the Samsung Gear VR [2]. This hardware, also called head-mounted displays (HMD), has seen great advances in the recent years and now users are able to use them wirelessly using a smartphone as the display. These advances led to an increased research in the area of virtual reality, which also led to an increased of interest in mid-air interaction.

Interaction techniques developed for IVEs aim to give users a suitable approach to object manipulation. These techniques can be divided in three phases: selection, transformation and release. In our work we focus mainly on the first phase. Most approaches to object manipulation are not focused on the interaction with objects that are out of arms-reach. When the virtual environment is larger than the room the users are in, reaching and selecting far away objects becomes a difficult task. This means that users have to rely on techniques designed to select these objects, or move closer to the object they intend to select.

When considering the task of urban planning in a virtual environment, there is a need to select objects that are not in the immediate vicinity of the user in order to manipulate them.

1.1 Challenge

The interaction with objects that are out of arms-reach still pose some problems. These techniques follow, in general, three different approaches: ray-casting , arm-extension and volume selection.

Ray-casting based techniques suffer from hand and tracker jitter, where a small hand movement can have a severe impact in the manipulation of the objects. When considering objects that are further away from the user, the jitter will have a greater impact, as a small hand oscillation can cause a sizable

[1] www.oculus.com
[2] http://www.samsung.com/global/galaxy/wearables/gear-vr/

position difference. On the other hand, arm-extension techniques provide users with a more accurate approach to object selection. But in turn, more movements are required which can lead to fatigue. Volume selection techniques can effectively deal with some problems that are presents in ray-casting and arm-extension techniques, but there is a lack of research when the objects are out of arms-reach.

Even though some approaches are capable of out-of-reach selection, the selection of objects that are very distant from the user is still a considerable challenge.

1.2 Objectives

We focused on the problems that selections techniques currently have, considering that users have to select an object in order to perform transformations on it. Our intention was not to solve the hand and tracker jitter problems, but to tackle the problems that ray-casting, arm-extension and volume based techniques currently have. With the work of this thesis we intended to answer the following research question:

Does the combination of progressive refinement and cone casting improve user performance in out-of-reach objects' selection tasks in VR?

1.3 Approach

We approached the problem, the selection of objects that are out arms-reach, by first studying the pertinent related work. Our main field of study were interaction techniques in mid-air, but relevant techniques in tabletop and large screen displays were also object of study, as the the valuable aspects of techniques developed for tabletops and large screen displays can be carried over to the mid-air paradigm. In our technique we use a cone as our selection volume and we explored the best methods of manipulating this volume, when using it for out-of-reach object selection.

Besides the main area of focus, we also investigated the pertinent work in VR travel and cybersickness. A study was also conducted in this area with the objective of lowering the duration of the time spent travelling in the environment. The results of the study were used as an argument for the choice of travel technique to use in association with our selection technique. Lastly, we carried out an evaluation with three selection techniques to ascertain which was fastest and more accurate when selecting objects out-of-reach.

1.4 Contributions

The principal objective of this work is to determine which techniques are more appropriate when selecting objects out-of-reach in immersive environments. Three approaches were studied on this subject, which were then tested with real users. To aid in the design of our technique, we conducted an evaluation regarding navigation in immersive environments. We also propose a taxonomy for object selection in immersive environments.

- **Assessment of techniques for object selection**

  Most solutions available for mid-air out-of-reach selection are still mainly based on three approaches: ray-casting , arm-extension and volume selection. We propose a technique for out-of-reach object selection, which we then validated by comparing it against two other approaches. Our objective was to understand if combining progressive refinement with arm-extension and cone casting would provide better results when selecting objects out-of-reach.

- **Evaluation of travel technique in VR**

  Travel in immersive environments and the effects of cybersickness are not very well understood. In order to design a better out-of-reach selection technique, we conducted an evaluation with the objective of reducing the duration of the time the user would be travelling and also the discomfort felt. We used the results from the test to justify the choice of travel technique that is used in our technique.

- **Taxonomy for selection techniques in IVEs**

  We introduce a taxonomy for object selection in immersive virtual environments. We classify techniques based on their three key characteristics: Reach, Cardinality and Strategy. This taxonomy allowed us to classify more in depth the techniques that were relevant for the design of our proposed solution.

1.5 Publications

The work that has been developed for this thesis has led to one publication which was evaluated by a panel of experts and was accepted in an international conference.

- Daniel Medeiros, **Eduardo Cordeiro**, Daniel Mendes, Maurício Sousa, Alberto Raposo, Alfredo Ferreira, Joaquim Jorge, *Effects of Speed and Transitions on Target-based Travel Techniques*, ACM 22nd Symposium on Virtual Reality Software and Technology, Munich, Germany, November 2016

- **Eduardo Cordeiro**, Daniel Medeiros, Daniel Mendes, Maurício Sousa, Alberto Raposo, Alfredo Ferreira, Joaquim Jorge, *Evaluation of Travel Techniques for Virtual Reality*, 23º Encontro Português de Computação Gráfica e Interação, Covilhã, Portugal, November 2016

1.6 Thesis Outline

In Chapter 2 we surveyed the current work in object interaction in virtual environments, as well as a describing the current work in the area of travel in virtual reality. We also introduce a taxonomy for object selection and a discussion of the relevant approaches in object selection.

Chapter 3 presents the proposed solution developed for out-of-reach object selection. In Chapter 4 we produce the prototype developed, introducing the architecture and setup chosen.

Chapter 5 details the user evaluations carried out, the first regarding travel in VR and the second to validate the developed solution when compared with two techniques from the literature. Finally, Chapter 6 provides a conclusion to the work and proposes the possible future work.

Chapter 2

Related Work

To understand how to better design a 3D interaction technique we reviewed the relevant related work in the area. Two major areas of the literature were the focus of our work, Object Selection in VEs and Travel in VR. We divided the object selection research in three distinct topics: Tabletop and Spherical Displays, Large Scale Displays and Fully Immersive. Based off the reach characteristic of the taxonomy we developed, and which is described later in this section, we subdivided the Fully Immersive approaches even further.

And finally, we analysed the relevant work in the area of Travel in VR, regarding navigation and cybersickness. Although our focus is not on how to travel inside a virtual environment, it was a necessary step in order to accomplish a well designed technique.

2.1 Object Selection in Virtual Environments

In the following sections we review the most relevant related work in selection of objects in virtual environments, that directly influenced the development of our technique. Even though our focus is in mid-air interaction techniques, there is a lot of research in adjacent areas that can prove useful. It is also helpful to determine what type of techniques have been proposed, and what their advantages and disadvantages are.

2.1.1 Tabletops and Spherical Displays

While interaction in tabletops and spherical displays is not the focus of our work, the techniques presented have some valuable aspects that can be translated into the mid-air paradigm.

Bezerianos et al. [1] developed the Vacuum, a remote reaching interaction technique that facilitates the selection of distant objects (Figure 2.1). Their technique relies on a circular widget with a user controllable arc of influence which the user can control using the pen/touch. Users can adjust the direction and angle of the influence arc dynamically during an operation. When an object is selected and vacuumed, a proxy is created near the user, which he can then manipulate.

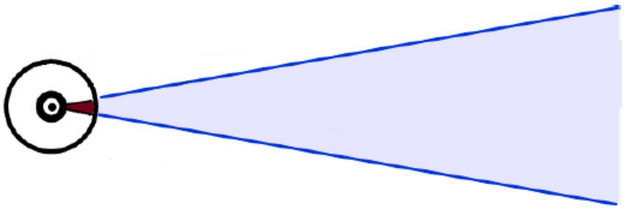
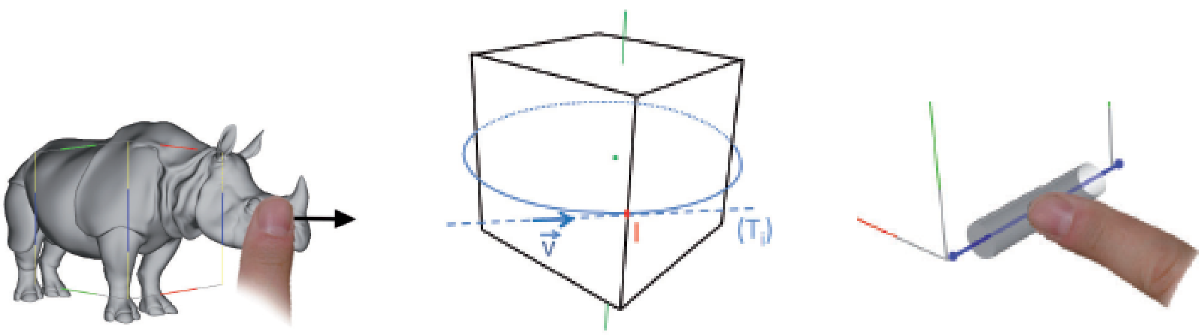


Figure 2.1: The Vacuum [1].

They compared the vacuum to drag-and-pop and the pick techniques [2]. The experiments indicate that vacuum perform similarly to direct picking and drag-and-pick, except for when the targets were located very far away. The Vacuum scales well to multiple target selection and performs significantly better in this scenario than existing techniques. The technique has a clear flaw relative to drag-and-pick, the scale of the proxies.

Cohé et al. [3] explored how widgets that are operated by mice and keyboard could be adapted to the tactile paradigm. They developed the tBox, a touch-based transformation widget that appears as an always visible wireframe box, and the transformations to the box are directly applied to the object (Figure 2.2). The widget favours direct and independent control of 9 degrees of freedom; 3 translations, 3 rotations and 3 scalings. In the informal study conducted, participants were asked to play with the interface in a scene composed of several objects that could be used to assemble a character, but with no precise goal. The participants approved the rotation mechanism and indicated it worked well, and so did the translation and scaling.

Intersecting 3D objects on a tabletop is not a trivial task, since the object is bound to the display surface. Hilliges et al. [5] created two different systems to tackle this problem. The first system is a rear-projection system where users can interact with the objects using standard 3DOF (modified SecondLight [6]). The goal of this technique was the way the feedback from shadows affects the overall experience of the user when interacting with a tabletop (Figure 2.3). Informal tests revealed that users commented that the hand shadows gave them a greater sense of interacting with the virtual objects, and some users thought the pinching gesture was not the most intuitive. The second system improves on the simulation of shadows by constructing a mesh from world coordinate values. They also propose a Grasping model



(a) Standard tBox transformation widget

(b) Rotation on the tBox widget

(c) Translation on the tBox widget

Figure 2.2: tBox widget and its transformations. [4]

(to compare to the pinching gesture), this technique allows 5DOF. They conclude by referring that, even though the second system was supposed to be an improvement on this first, both system have their strengths and weaknesses.

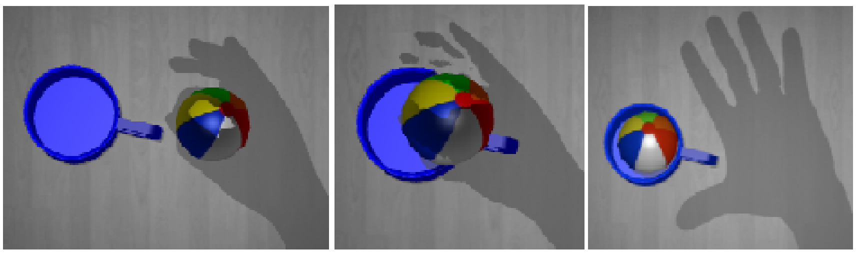


Figure 2.3: Virtual shadows cast by the users hands [5].

Bannerjee et al. [7] created Pointable, an in-air asymmetric bi-manual single-cursor manipulation technique to select and move digital content over interactive tabletops. The dominant hand points and acquires remote targets (Figure 2.4), while the non-dominant hand scales and rotates the target without the need to drag the target closer. Its design aimed to be an addiction to direct-touch, had low invocation and dismissal overhead, should allow in-place manipulation of remote objects, minimise physical contact and also be unobtrusive.

They performed three experiments to see if their technique could satisfy the previous goals. The first experiment aimed to compare the throughput of perspective-based pointing to touch; in the second experiment they wanted to compare the performance of multi-touch to Pointable in a standard translate/resize task; in the last experiment their focus was on observing the behaviour of the participants when they were free to choose their interaction. The results of their experiments suggest that Pointable can be used in conjunction with direct-touch, which was their main goal.

De Haan et al. [8] developed a technique for object selection in tabletop environments. It is a selection-by-volume technique that extends the ray-casting technique (Figure 2.5). It uses a scoring function that scores each object that falls within the conic volume with a value between 0 and 1. This rating is given based on how close the centre is to the pointing direction. The scoring of these objects is accumulated, and from this they managed to create a dynamic, time-dependent object ranking. The user can tell which objects has the highest score, or the active one, as the ray bends towards it. Their tests point out that IntenSelect improves the selection performance when compared to the ray-casting

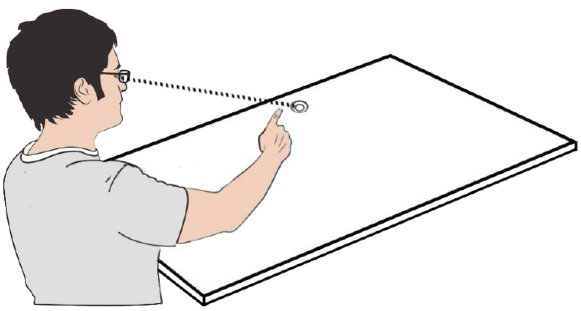


Figure 2.4: Perspective based pointing technique, Pointable [7].

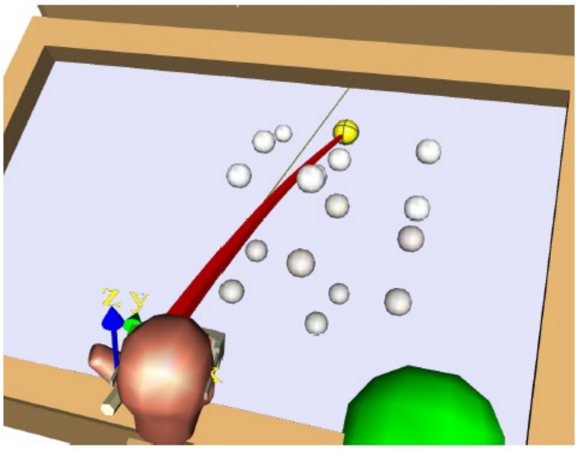


Figure 2.5: The IntenSelect technique [8].

technique. They also propose a time-dependent object ranking system, and that it proves notably helpful when objects are moving, occluded and/or cluttered.

Rapid aimed target pointing in two dimensions is a fundamental task in user interfaces, and has been studied heavily. Three-dimensional pointing on the other hand is far less well understood. The pointing task itself is more complex, as the cursor is usually controlled by at least 3 degrees of freedom. Teather and Stuerzlinger [9] investigated means of comparing 2D and 3D pointing techniques in stereoscopic screens with the aid of stereographic *CrystalEyes* shutter glasses and an emitter.

Five techniques were implemented and tested against each other (Figure 2.6). The Pen Touch (PT) use a tracked stylus and displays a 1 mm diameter cursor co-located with the stylus tip. Pen Ray (PR) also uses the stylus, but casts a ray from its tip into the scene. Mouse Cursor (MC) uses a mouse controlled 3D cursor that moves in the screen plane. Floating Cursor (FC) and Sliding Cursor (SC) are designed to address the mouse cursor occlusion issue for targets above the display.

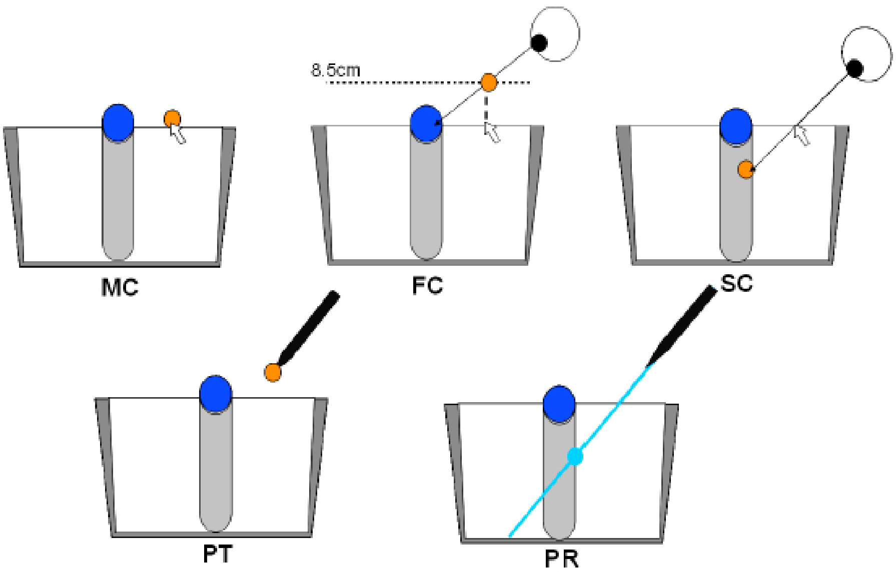


Figure 2.6: Top Left: Mouse Cursor. Top Center: Floating Cursor. Top Right: Sliding Cursor. Bottom Left: Pen Touch. Bottom Right: Pen Ray [9].

To compare these techniques they designed three experiments. They concluded that 3D pointing performance degrades when targets are displayed stereoscopically for 3D techniques, but not for 2D techniques. Simply using the Euclidean 3D distance rather than 2D distance into account seemed to predict 3D motions sufficiently well.

To study which approaches to 3D object manipulation were the best, Mendes et al. [10] compared four mid-air techniques in a stereoscopic tabletop environment. The 6-DOF technique lets users grab the object with one hand and dragging the object in space moves it in the three dimensions while wrist rotations control the object rotation. The 3-DOF Hand technique separates translations and rotations. Translations are done using the hand that grabbed the object and rotations are done by grabbing the space and keeping the other hand on the object. They implement the Handle bar technique [11], where the middle point between both hands is used to manipulate the object. In the Air TRS technique, the first hand that grabs the object moves it, while the other hand rotates and scales the object. They used non-intrusive tracking of the head and hands, which improved perception and allowed walk-up-and-use and mid-air interactions. The 6-DOF Hand approach was favoured by the participants of their user tests, as it reproduced direct interactions with physical objects, even tough the Handle Bar technique was considered as fast as the 6-DOF Hand.

Grossman et al [12] present an exploration into 3D selection in volumetric displays. In a first experiment they found that a ray cursor technique was preferable to a 3D point cursor when working in single target environments 2.7. Four new techniques were then designed and implemented that focus on the difficulties related to dense target environments. The *depth ray* technique had the best performance, and significantly reduced movement time, error rate and input device footprint when compared to the original 3D point cursor. The authors also show that these techniques can be applied to other interaction paradigms.

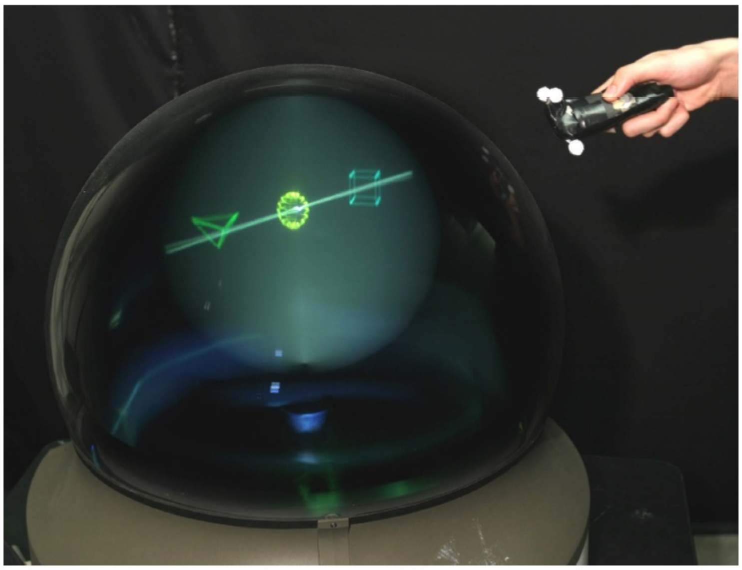


Figure 2.7: Ray cursor selection [12].

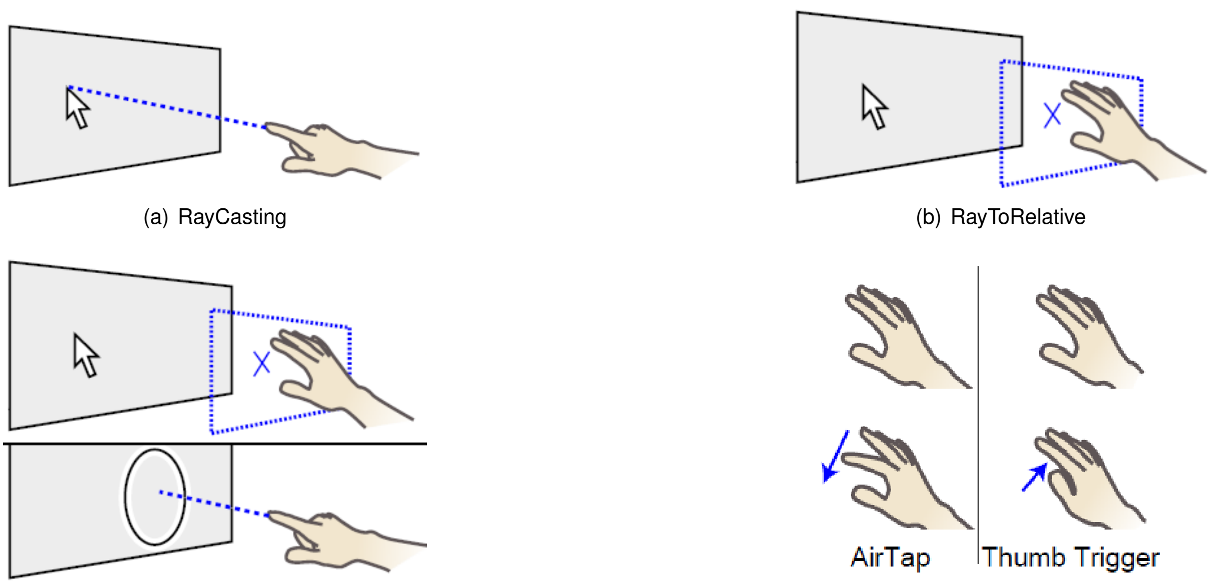
2.1.2 Large Scale Displays

Large scale displays have an inherit problem, users interact with large 2D surfaces to manipulate objects that are not in that 2D plane. Techniques used in large scale displays can be useful to our work as some objects can be placed away from the user, making it hard to select and manipulate them.

Vogel et al. [13] created and evaluated three techniques for gestural pointing and two for clicking. They developed three Clicking techniques and two Ray Casting techniques (Figure 2.8) based on five characteristics: accuracy, acquisition speed, pointing and selection speed, comfortable use and smooth transition. They enhanced these techniques with auditory and visual feedback to compensate for the lack of kinesthetic feedback that is present in freehand interaction.

They found that although Ray Casting's error rate prevented it from being a practical technique, no major significant effect between the Relative and RayToRelative techniques in terms of selection time or error rate was found. Both the relative pointing techniques behaved different in the trials, and they state that the number of recalibrations activations and recalibration times were the defining factors. Even though the longer recalibration times of the RayToRelative technique, it did not impact its overall selection time.

Rotate-Scale-Translate (RST) interactions have become the *de facto* standard when interacting with two-dimensional contexts in single and multi-touch environments. Reisman et al. [14] extended these principles to the 3 dimensions. They presented a screen-space method that allows direct control in 2D and 3D on a multi-touch surface, where multiple points of contact are at the user's fingertips (Figure 2.9). In their initial tests, they state that when controlling 2D planar objects (in 3D) they could not differentiate between their manipulator and other 3D RST controllers. They use 3 fingers to perform rotations (adding 2 DOFs). They also introduce a way to constrain an object's motion, penalties. They are calculated



(a) RayCasting

(b) RayToRelative

(c) Hybrid RayToRelative

(d) Clicking Techniques

Figure 2.8: Distant pointing techniques [13].

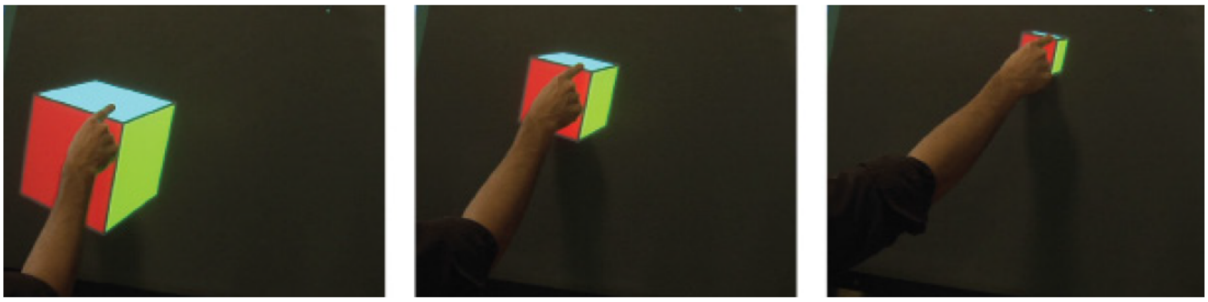
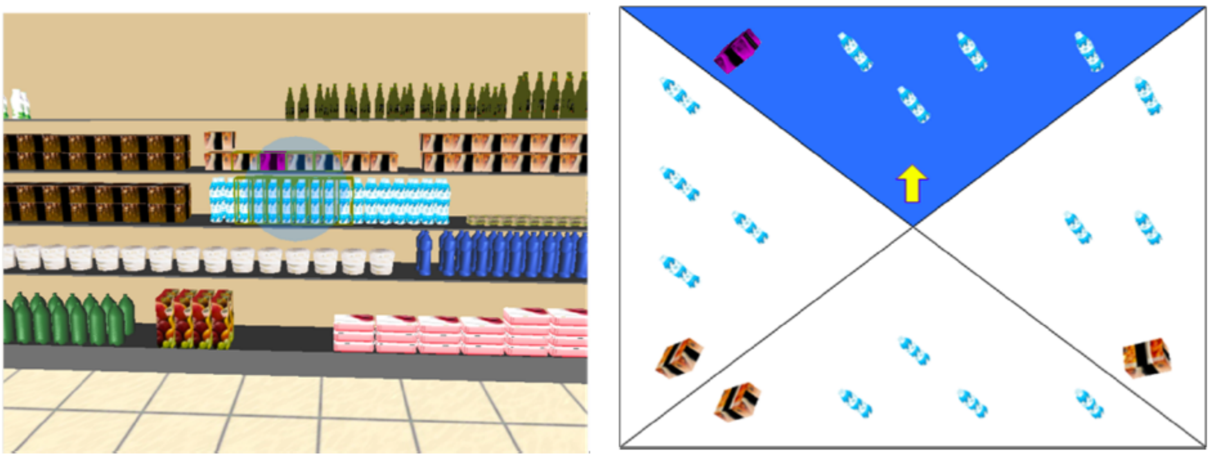


Figure 2.9: Control over a cube in a large scale display [14].

differently for the various types of interaction, and since these penalties can be weighed they can assign a harsher penalty to some motions than others.

Argelaguet and Andujar [15] studied the feedback provided by techniques for accurate pointing on stereo displays. Most techniques used in ray-casting provide inaccurate feedback regarding the alignment of the tool used for pointing and the target. They also report that they perform unsatisfactory when there is a need for accurate pointing, even when the scenarios had low levels of occlusion. They found that controlling the direction the user is pointing with hand orientation results in better performance than with the hand position. Their technique, viewfinder, overcame the main limitations of techniques that were studied and tested. Their work provides three notables advantages. Firstly, a much smaller impact on the immersive view. It is also insensitive to poor stereo separation and ghosting effects, which are still present in monocular vision And lastly, it allows non-tracked users to clearly identify the object pointed to by the tracked user.

Kopper et al. [16] developed progressive refinement in a attempt to solve an inherent problem in ray-casting selection techniques in large scale displays: hand and tracker jitter. They designed a progressive refinement technique called Sphere-casting refined by QUAD-menu (SQUAD). This technique works in incremental phases. The first phase uses a modified version of ray-casting where a sphere is cast onto the nearest intersecting surface to determine which objects are selectable (sphere-casting). After the



(a) Sphere-casting

(b) Quad-menu

Figure 2.10: Progressive Refinement technique [16].

Figure 2.11: Continuous Zoom technique [17].

first phase, all object that were inside or touching the sphere are distributed among four quadrants on the screen (Figure 2.10 (a)). A user can refine the selection process by repeatedly pointing anywhere in the quadrant that contains the object they are searching for (Figure 2.10 (b)).

In their experiments they compared SQUAD to standard ray-casting. There was a performance trade off between immediate techniques that use one precise action and progressive refinement techniques that require low precision at the expense of several steps. SQUAD is significantly faster for selection of small objects or selection in low-density environments. SQUAD's task completion time grew linearly as the number of refinements increased, as opposed to standard ray-casting as time increased is exponential as targets become smaller. Progressive refinement techniques show very high levels of accuracy because precise pointing is not required, which is a major cause of hand jitter.

The work of Kopper et al. [16] on progressive refinement technique was extended by two distinct authors. Bacim et al. [17] designed two new techniques based on progressive refinement and compared them against the standard ray-casting technique. For tasks that involved visually small targets, the SQUAD, Discrete Zoom and Continuous Zoom (Figure 2.11) techniques performed similarly or better in terms of speed as ray-casting. They also note some design flaws of the SQUAD technique, where the sphere-casting is not a good fit when the objects are distributed in depth through the environment. The results from their tests indicate that the multiple easy selection steps can be faster than a single difficult selection, and that in terms of accuracy, the progressive refinement techniques were clearly better.

Two more techniques were developed based on ray-casting and SQUAD [18]. The results indicate that the performance of the new techniques, Zoom and Expand, is significantly higher than the standard ray-casting and SQUAD techniques. The Zoom methods brought the user closer to the desired objects, but still did not solve the occlusion problem. In the Expand technique, the quadrant arrangement of SQUAD was redefined, by using a virtual grid that filled with screen with the selected objects. The level of density of the scenarios played a big role in the accuracy and speed of each technique. The Expand technique was also mentioned as the best over the course of different scenarios.

2.1.3 Fully Immersive Displays

A great deal of research has already been done in techniques used in immersive environments. We categorise these techniques largely based on one of the key aspects of our taxonomy for object selection, the reach of the technique, as described in Section 2.4. Arms reach techniques can be here considered as manipulation techniques, as the user simply extends his arm to the object and then can perform more complex operations on it.

Techniques intended to be used in arms reach are usually manipulation techniques, and the selection type used is usually direct. Improved reach techniques are those where the user can reach farther into the environment than the length of his arm but still have limits. Lastly, infinite reaching techniques are able to select any object in the environment.

Arms Reach

Understanding the importance of degrees of freedom (DOF) in 3D manipulation is important, since manipulation in mid-air demands the use of all 6 DOF's, even though some constraints can be applied.

To study the impact of DOF separation and integration on the user's performance, Veit et al. [19] compared two techniques, Indirect Rotations (IR) and Bi-manual Plane Constrained Rotations (BPCR), on the manipulation of the DOF's of a 3D orientation task (Figure 2.12). In the IR technique the user grabs a virtual manipulator positioned immediately at arms-reach and orientates it just by rotating the hand. The BPCR technique consists of commanding a rotation around an axis of the objects coordinate system by moving his finger along the corresponding screen axis. Users can use both hands simultaneously or successively to perform rotations by touching the screen. The dominant hand gives users access to two specific axes of rotation and the non-dominant hand gives him access to the third axis.

In the results observed in the user tests two of their hypothesis failed. The IR technique was less precise than the BPCR technique and that for small orientation and complex operations BPCR is more efficient than IR. Using MDS they concluded that using the BPCR technique users decompose one 3 DOF orientation task into three 1 DOF sub-tasks. IR and BCPR showed the same level of precision, but BCPR was the fastest technique.

Techniques that focus solely on transformation are also importation to our work, as we need a solution that allows for both selection and transformation of objects. Wang et al. [20] created a markerless hand-tracking system using depth cameras, capable of allowing 6 DOF manipulation. The system relies on a small set of gestures that are comfortable, precise and easy to remember. They aimed to correctly identify pinching gestures without markers, for which they reported a 99% accuracy. The technique has

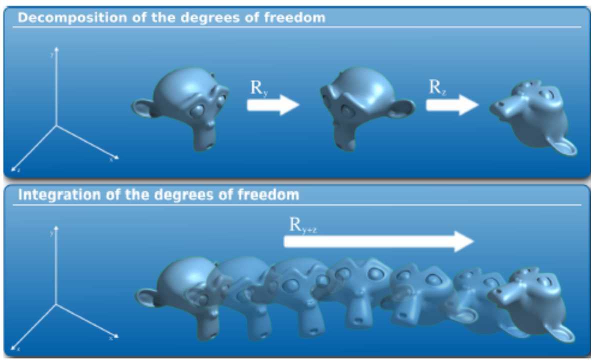


Figure 2.12: Two strategies to achieve an orientation task [19].

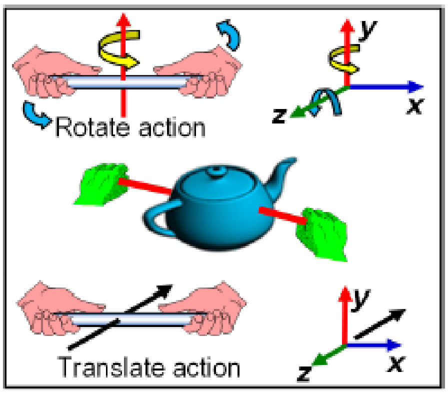


Figure 2.13: Operations available in the handle bar technique [11].

some issues, since it requires the hands to be segmented from the background with skin-tone detection and the use of a long-sleeved top because the user's skin-toned arms are not modelled.

A novel approach to mid-air interaction was the Handle bar technique by Song et al. [11]. It used three gestures to achieve new forms of interaction and manipulation: POINT, where the hand is closed and only the index finger is pointing forwards, OPEN where the hand is completely flattened and open and CLOSE where a full fist action is performed (Figure 2.13).

The handle bar technique enabled the users manipulation of multiple objects at the same time. These manipulations are familiar, as the translation and scale of the objects are done as if they were a single object, and the rotations are executed in the mid-point of the handle bar. Aligning multiple objects lets users to bundle several objects in the same handle bar, while tilting the handle bar allowed for a quicker alignment, as gravity is involved. There were still some problems with this technique, as some users complained of arm fatigue after 20-30 minutes of usage. There was also slight handle bar wobble that would occur during the hand gesture change and inter-hand occlusion could occur during the rotations.

Improved Reach

The Go-Go technique [21] uses a non-linear mapping function to map the users arm movement to the virtual arm (Figure 2.14). This technique uses a local area around the user and while the users hand is in the area the virtual hand behaves as normal, and when it goes beyond the area limit, the virtual hand begins to move outwards faster. Although no translation was offered, users could select, rotate and scale the objects.

Auterii et al. [22] integrated the Go-Go and PRISM techniques to increase precision in 3D manipulation. They tackle the problems that are present in previous techniques such as ray-casting and HOMER. They apply the PRISM technique to Go-Go, since it is a technique that resembles PRISM as both adjust the control display (C/D) ratio based on heuristics. To combine both techniques and to achieve the desired cursor position they calculated the C/D ratio by combining the ratios of both techniques, in the form of a product of the two C/D ratios. A note is made in regarding to a limitation of their design. If a user extends their hand out slowly, the effects of Go-Go are offset by PRISM scaling, and reach extension is limited.

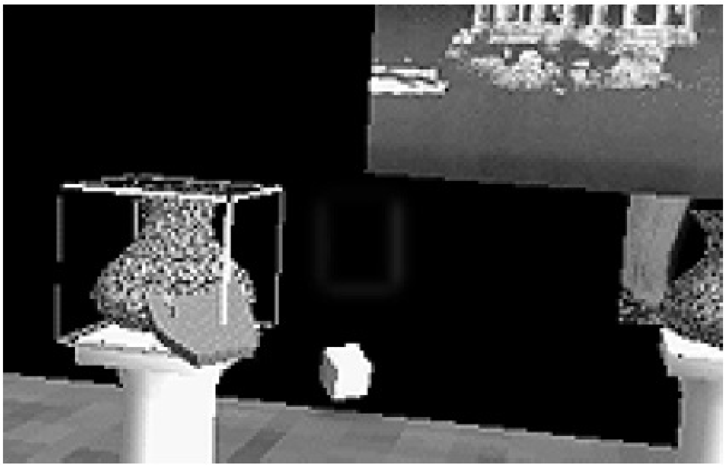


Figure 2.14: The Go-Go interaction technique [21].

The effect of their technique is to deliver similar levels of perceived precision at all distances, although it does not permit users the same actual level of precision for all distances. The tests with users report that when using Go-Go + PRISM the average numbers of completions was substantially highers than with Go-Go alone.

Stoakley et al. [23] exploration into World in Miniature used a scaled down copy of the world the users were immersed in, which allowed for direct manipulation with the objects in the actual world (Figure 2.15). Even though this was a suitable technique, it faced some limitations. The WiM does not allow for precise object manipulation, since a small movement in the scaled down world moves the actual object more than intended. They also revealed that it was a very exhausting technique, since users had to keep their non-dominant hand lifted throughout out the experiment.

The Voodoo Dolls technique [24] is a two-handed interaction technique for manipulating objects at a distance in immersive virtual environments. It offers the users a way to work seamlessly at multiple scales, allows both visible and occluded objects to be manipulated. The dolls are hand held, short-term

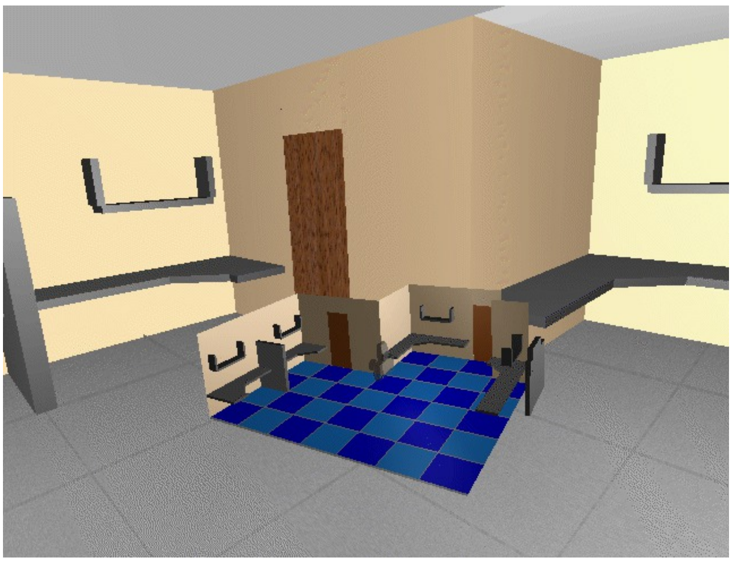


Figure 2.15: World in Miniature [23].

Figure 2.16: Creating a doll in the Voodoo Dolls technique [24].

objects: users create dolls to represent specific objects, and the dolls vanish when they are no longer needed. The dolls are small copies of the objects they represent and are used in pairs, one in each hand, and serve two different but complementary purposes depending on the hand holding them. Changing a doll from the right hand to the left, changes the dolls' mode, and can be used as a frame of a reference to manipulate another doll.

Dolls are created when the user selects an object using an image plane technique [25] as depicted in Figure 2.16. Selecting an object creates a doll that is scaled to a convenient working size and placed in the user's hand. As a user holds a doll in his left hand, the system created dolls for nearby objects to provide context for the interaction. It allows manipulation over objects that are occluded or are too small on his image plane.

Even tough most mid-air techniques take place on tabletops or are aided by an head-tracking system, Benko et al. [26] implemented a gesture-based interactive experience in an unusual surface, the dome. They chose pinching as their basic unit of interaction, mapping different actions to different types of pinches. Additionally gesture-invoked speech recognition was incorporated (gesture + speech input) as a mean to navigate between the various visualisations. A retro-reflective tape on the users' hand was the method used for mid-air clicks, while performing an "hand clasp" gesture. Users also used a laser pointer as a mean to point at a specific area of the dome, but was also observed that most users used their hands shadows as the remote pointer. Users mentioned that the gestures were neither self-evident nor easy to learn.

The lack of haptic feedback, limited input information, limited precision and a lack of a unifying framework for interaction are some of the problems of Virtual Reality manipulation. Mine et al. [27] explored new means of compensating for the lack of feedback in IVE's. Proprioception, as described by the authors, is a person's sense of the position and orientation of his body and limbs. They consider that working in arm's reach gives the user's natural advantages: uses proprioceptive information; provides a more direct mapping between hand motion and object motion; yields stronger stereopsis(perception of depth and 3-dimensional structure obtained from both eyes) and head-motion parallax cues and provides finer angular precision of motion.

Two formal studies were done where the conclude that the lack of physical work surfaces and haptic feedback makes the controlled manipulation of VR objects more difficult. Humans rely on naturally

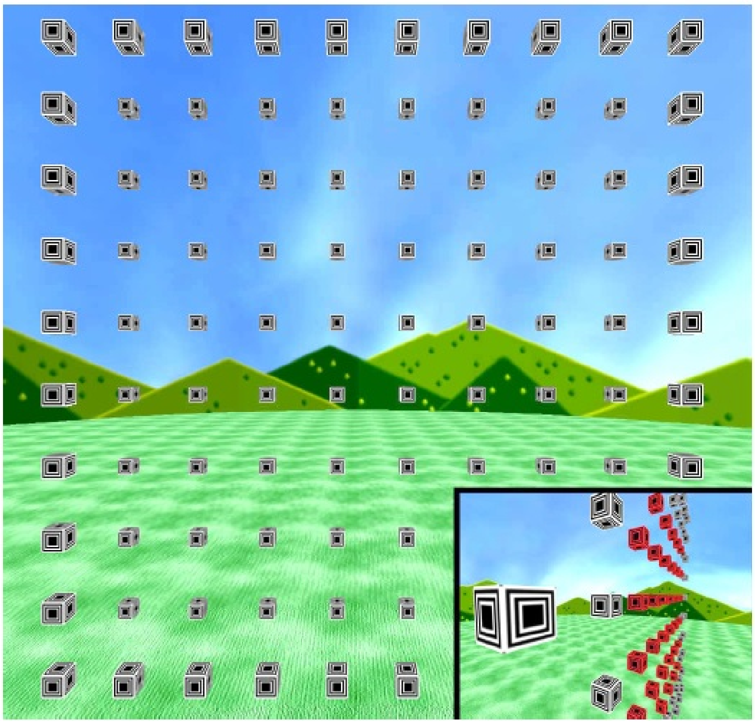


Figure 2.17: Summative Experiment Environment used [28].

occurring physical constraints to help with the motion of objects. Users in a virtual world must typically do without the fingertip control they rely on for the fine-grained manipulation of objects in the real world.

Multiple object selection (MOS) in immersive Virtual environments has not been the target of comprehensive research. Lucas et al. [28] propose a taxonomy for MOS selection and also describe four techniques for selection of multiple objects. Two of their techniques are serial (only one object can be marked per operation) and the other two are parallel (one or more objects can be marked by operation). In addition to this, they also considered two metaphors for interaction: 3D spacial metaphor and the pen and tablet metaphor. From the two studies conducted (Figure 2.17), the first revealed that serial techniques were easier to understand and that they required less initial setup time. The parallel techniques usability suffered in comparison, as they say that user strategy and sophistication played a bigger role when using these techniques. The second study showed that parallel MOS techniques can be superior to serial MOS techniques under certain conditions. They also question if parallel techniques should be used in addition to serial ones.

Stenholt [29] notes that the difficulty of efficient MOS increases with the number of objects that need to be selected. When talking about small-scale MOS (selection of few objects), serialising existing single-object selection techniques can be favourable. They explored various MOS tools for large-scale MOS (when the number of objects to be selected are in the hundreds or thousands). Three techniques for 3D MOS were tested, the brush and lasso from the literature [30] and a new technique called called the magic wand. Using the 3D magic wand was very fast, but also very sensitive to the geometric scenario which made it very easy to use or absolutely impractical. They state that the 3D spherical brush is a good candidate for a simple, general 3D MOS tool that is applicable to many scenarios. The study indicated that the performance of MOS techniques depends heavily on the geometric scenario.

Infinite Reach

Bowman et al. [4] evaluated a series of techniques for manipulating objects in IVE's. They separated the techniques in two categories: arm-extensions and ray-casting (Figure 2.18). The arm-extension techniques make use of the virtual arm to select and manipulate the objects, and they are the Go-Go [21], the Stretch Go-Go (a version of Go-Go where the arm can stretch to the infinite) and an indirect stretching technique where users press buttons on a 3D mouse instead of using arm motions. In regards to ray-casting techniques, they evaluated the Fishing Reel, where a user, after selecting an object via ray-casting, can bring it closer or farther using mouse buttons. And lastly the HOMER technique where a user grabs an object using a light ray, but instead of the object becoming attached to the ray, the virtual hand moves to the object position. The HOMER interaction techniques takes the best aspects of the tested techniques and combine them in a seamless way to maximise ease of use and efficiency.

Wilkes et al. [31] continued and expanded on Bowman's HOMER technique, creating the Scaled HOMER. This technique allows precise manipulation at both near and far distances based on velocity-based scaling. By adding velocity-based scaling, they provide more precise object control. So, when the hand is moving quickly, the scaled velocity will be equal or greater than the actual velocity, and when the hand is moving slowly, the scaled version will be less than the actual velocity. They developed this technique since HOMER had several limitations: users reach was limited by the length of their arm and there was not any support for precise positioning at a distance. Overall, scaled HOMER offered a significant improvement over the basic HOMER technique, and they report that the mean task time for HOMER was almost double the task time for the Scaled HOMER, which indicates that velocity-based scaling has a benefit in a wide range of task conditions.

Pierce et al. [25] presented a group of techniques that use the 2D image plane concept for selection, manipulation and navigation in virtual environments. The Head Crusher technique, where a user places two fingers above and below the object they want to select, the Sticky Finger, where one finger is placed

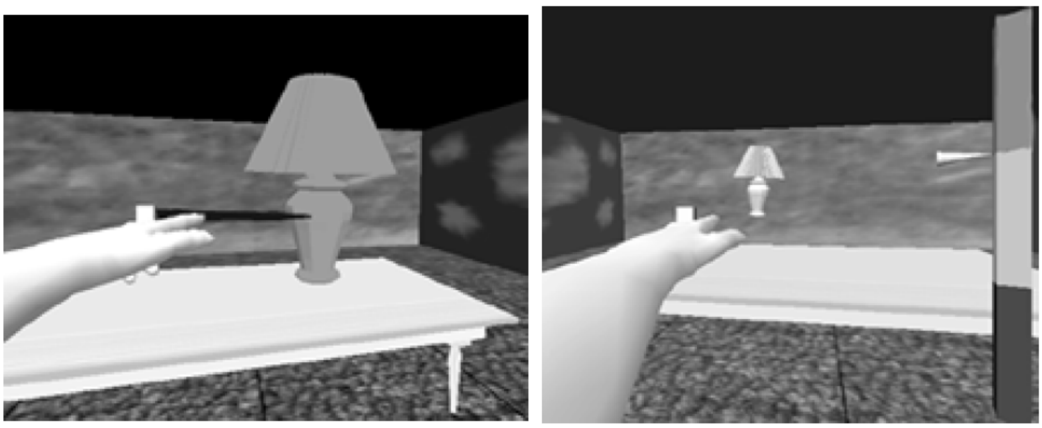


Figure 2.18: Object selection techniques [4].

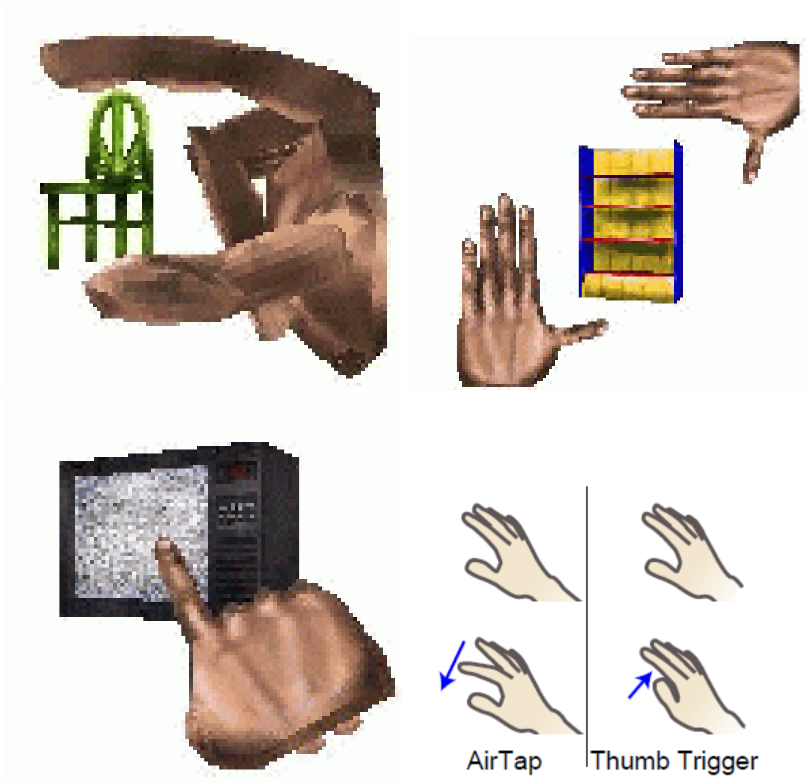


Figure 2.19: 2D Image Plane techniques [25].

over the object and a ray is cast to it, Lifting Palm, where one hand is placed flattened just below the object, and finally Framing Hands, where the users use both hands to "frame" an object (Figure 2.19).

Their techniques can also be used for navigation within the environment. The distance of the user from the object can be held constant to let the user orbit the object. As an alternative, they can vary the user's distance from the object as a function of the user's hand motion. This was achieved using a linear function that changes the distance the user is from the object using the changes in the hand's distance. The user's could also scale down the world and use translations over it to navigate.

They completed informal tests with six users to establish if people had problems using the techniques, to which they concluded that there were three major problems. The first is choosing the left or right eye's image when the scene is rendered in stereo. They solved this by rendering the scene monocularly. The second problem detected was arm fatigue from having to work with the arms constantly extended. Their techniques attend to this problem by allowing the user to move the object to a more comfortable position. The hands were rendered translucently so that the objects could be seen through them, to offset the users hands obscuring the objects.

Liang et al. [32] developed the Spotlight technique as means to select objects without using a mathematical ray, using a cone instead. Objects inside the cone are candidates for selection, but the object closer to the centre line of the cone should be selected. They state that the selection of objects that are farther away from the user are easier to select, as users do not have aim to perfectly. The selection of objects that are close is as easy as with using a ray. A transparent cone was drawn to indicate to the

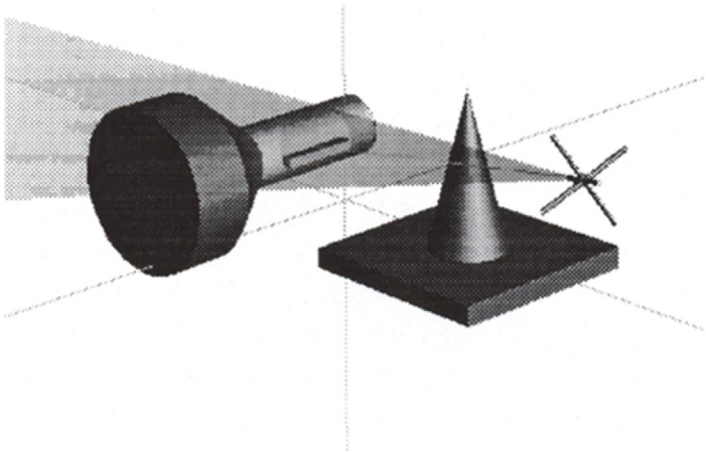


Figure 2.20: Spotlight Selection [32].

users which objects were being highlighted, and the light would be more intense on the ones that were closer to the centre line (Figure 2.20).

Forsberg et al. [33] developed two techniques for selection in VEs, called *aperture* and *orientation*. The aperture technique is a modification of the spotlight technique and it consists of casting a cone from the participant's dominant eye to the tracker's location (represented as a cursor in the VE). The volume of the selection cone is given by the distance between the eye and the cursor, and can be adjusted by moving the hand in or out (Figure 2.21). The orientation technique augments the aperture technique, by adding an orientation element to the selection process. Objects that are in the selection volume are matched with the orientation of the tracker, and the one with the closest orientation is chosen. They also state that the aperture technique provides three advantages in comparison to the spotlight technique: it closely resembles how we point at distant objects, incorporates volume selection as the spotlight technique [32] to lower the effects of tracker and user-induced noise and since the only visual presence is the aperture cursor, the screen is less cluttered.

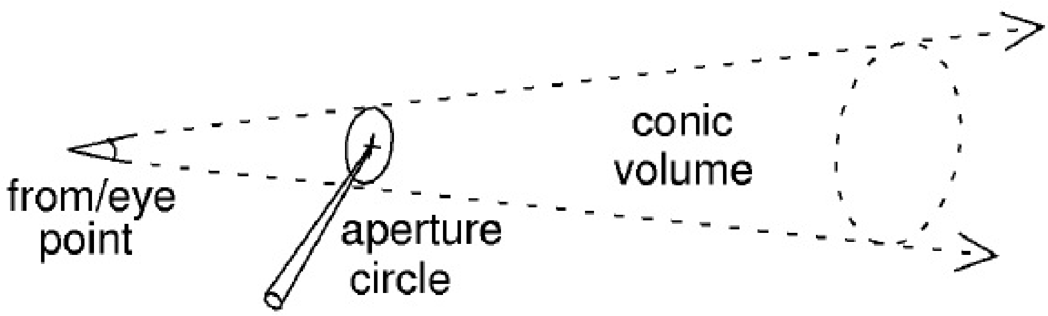


Figure 2.21: The conic volume of the aperture selection technique. [33]

2.2 Travel in Virtual Reality

We surveyed ways of moving a user in the virtual world, and the section that follows relates to two interconnected domains of previous research, travel techniques and the causes of cybersickness in VR environments. We review this area of the literature and it is very relevant to our proposed technique for out-of-reach object selection in VR. Another crucial issue when considering travel technique design is cybersickness. It can be a consequence of spending too much time of being inside an immersive environment or travelling inside it.

Travel is the most common interaction in VR applications. Travel is the simple act of the user moving to a position where they can perform a more important task and is composed by three main parts: Direction/Target Selection, Velocity/Acceleration Selection and Input Conditions [34]. It differs from the cognitive part of navigation, which consist in the planning of the path taken. It is important when travelling in VEs that the user establish a sense of presence within them. Bowman et al. [35] proposes a list of quality factors for a efficient travel techniques which are: Speed, Accuracy, Spatial Awareness, Ease of Learning, Ease of Use, Information Gathering and Presence.

Travel techniques could also be divided in two subcategories, on *Explore* tasks the user moves freely on the VE without a predetermined goal and *Search*, where he has to reach a specific checkpoint. The more natural the technique, the more efficient it is on travelling tasks on VEs, specially on Explore tasks. Suma et al. [36] proves this by comparing real walking with Steering techniques [37] such as translation based on torso (torso-based), pointing (pointing-based) and gaze direction (gaze-directed). But in some cases, real walking is not desirable because of physical constraints of the indoor space available and obstacles on the physical world. Also, on typical Search tasks, the physical effort of reaching the goal can also be inconvenient and could lead to physical fatigue.

One of the factors that makes navigation difficult in virtual environments is user disorientation. Smith et al. [38] considers that the two leading causes are the absence of visual cues and problems with navigating too close to or through objects. Another factor that could lead to disorientation on VEs is the lack of control while travelling. Travel techniques where the users do not physically translate their bodies, such as Steering Techniques, can allow the maintenance of a satisfactory spatial awareness. This is justified by having some control over the movement of the user's head to explore the VE [39]. Even though, Target-based techniques (where the system shifts the user's position automatically based on predetermined speed parameters) may be more appropriate for users who are more susceptible to simulator sickness in comparison with Steering techniques [40].

Cybersickness can be distinguished from motion sickness, in that the user is in a stationary position but has a sense of self-motion through the moving environment. LaViola Jr et al. [42] suggest that there are too many causes for cybersickness and that there is no reliable approach for eliminating this problem. They also note that all theories regarding the causes of cybersickness should take the individual into account.

Lin et al. [43] conducted a study that concluded that VR-based dynamic environments can physiologically and realistically cause motion sickness on its users. They describe that these phenomena were

(a) Parrot roller coaster

(b) Helix roller coaster

Figure 2.22: Roller coaster comparison [41].

almost the same as those induced in a real environment. When comparing the onset of cybersickness on users of two distinct roller coaster experiences [41], the one with higher level of graphic realism was considered to be the one that caused a faster onset of nausea (Figure 2.22).

So et al. [44] investigated the effects of navigation speed on the level of motion sickness with head-steered techniques in immersive virtual environments. They report that nausea and vection (illusion of self-motion in the opposite direction caused by wide field of view in the absence of physical motion) sensation increased with the raise in navigation speed from $3m/s$ until $10m/s$ where they stabilised until the $59m/s$. They conclude that the mean offset time of vection was earlier than that of nausea, which they say is consistent with the knowledge that cybersickness is a type of vection-induced motion sickness.

Fernandes and Feiner [45] refer to VR sickness as a direct factor from the diminished field of view (FOV) present in the head-mounted displays available. To examine this effect, they dynamically change a seated user FOV while they explore a VE. Even with a small number of participants, the FOV restrictors helped the users feel more comfortable and able to stay in the VE for longer. These restrictors were not noticed by the majority of the users.

2.3 Taxonomy

Before making a comparison between the techniques that we found most relevant in the literature, we describe a taxonomy for the classification of selection techniques. We created this taxonomy based on Bowmans taxonomy for object selection, transformation and release [46]. As previously mentioned, we will not be focusing on either the transformation and release aspects of interaction technique, as this is not the focus of our work.

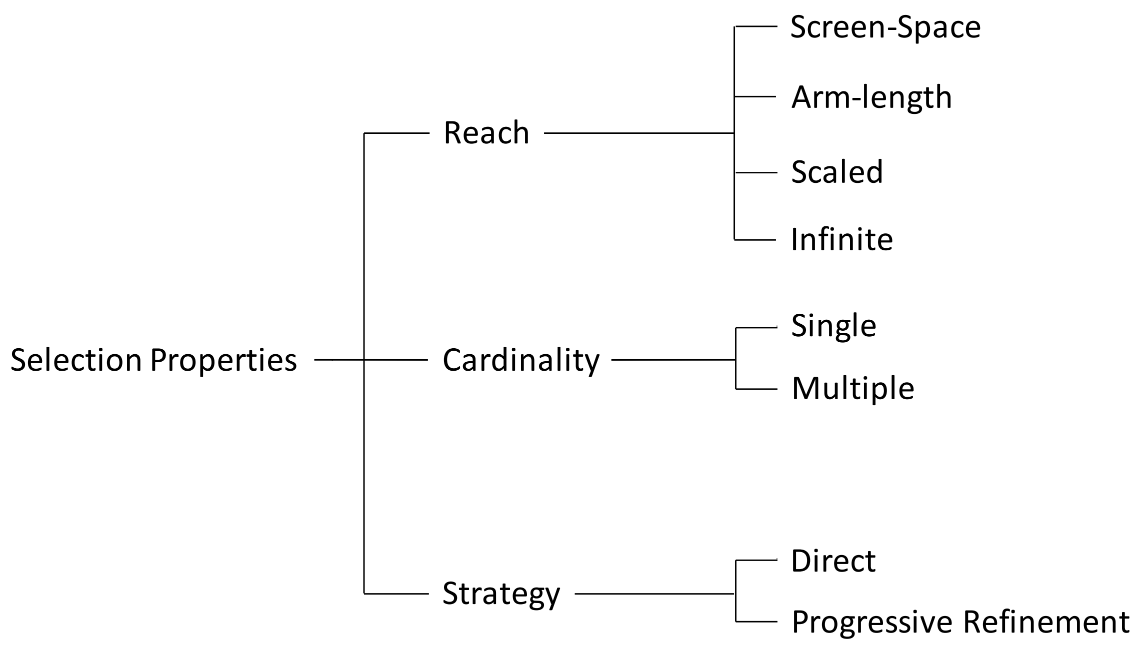


Figure 2.23: Taxonomy for object selection techniques.

In this taxonomy we distinguish each technique based on three different criteria (Diagram 2.23). First we classify each technique based on its Reach, which represents how far an object can be selected, and is divided in four groups. Techniques where the environment is either 2D or 3D and displayed in a desktop context or in a large scale display is classified as screen-space. Arm-length refers to the techniques where the length of the users arm is limit to where it can reach. Techniques that are classified by Scaled are the ones where the extent of the users reach is greater than its arms length, but can not reach infinite. Infinite reaching techniques are the ones where there is no limit to where the user can reach an object.

Next we classify the techniques in regards to their Cardinality. The techniques that can only select one object per operation are classified as Single. The other techniques, that are capable of selection multiple objects on a single operation, are referred as Multiple. Techniques capable of selecting multiple objects can be then decomposed into Serial and Parallel [28]. We are only focused on selecting one object per task, and to that effect we will not classify the techniques based on their criteria.

Lastly, we distinguish between the Strategy adopted when selecting objects. Either the technique only allows selection using a Direct method, where only one step is completed to select the object. This means that the user only has one attempt to select the correct object. Or on the other hand, a Progressive Refinement metaphor, where several steps are involved when selecting the object. These steps involve the users selecting a group of objects and then refining this group, by selecting fewer objects, until only one object is selected.

Besides offering a taxonomy on techniques focused on object selection we also classify the environments of these techniques (Diagram 2.24). We create a simplified version of the taxonomy introduced by Grossman et al. [47] for Display Properties. We classify only the Perceived Space and the Input Space of the VE of the techniques. The perceived space refers to the possible spatial locations where

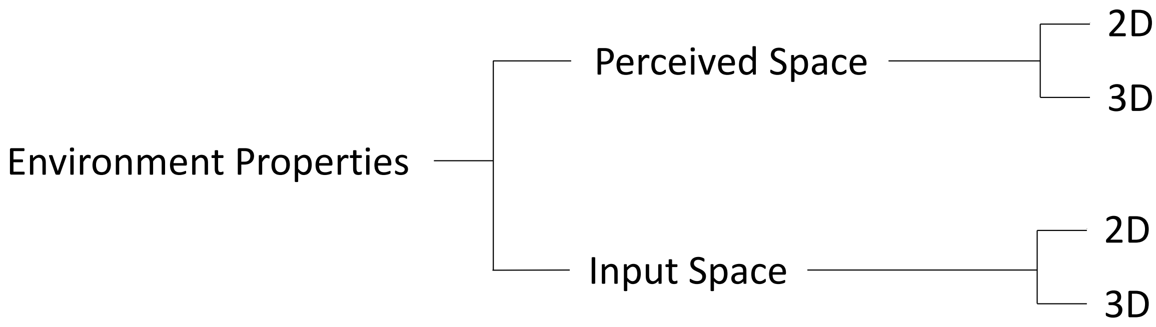


Figure 2.24: Classification of the environment of selection techniques.

the displayed image exists based on stereoscopic depth cues. Techniques that work in a traditional 2D tabletop, desktop or large scale display we consider them as having a 2D perceived space. When using an HMD, shutter glasses or a volumetric display, we categorise these techniques perceived space as 3D. Input space is the physical location where the users provides input. In regards to the input space of the techniques, we classify techniques as 2D when they are used on large scale displays or when used directly in 2D tabletops (touching the surface). Techniques used in mid-air are considered 3D.

2.4 Discussion

After presenting the relevant related work, we now make a comparison between the techniques that we found most significant to our proposed solution, as well as their advantages and disadvantages. Table 2.1 shows a comparison between techniques for object selection, based on the taxonomy described above.

The initial concept of progressive refinement introduced by Kopper et al. [16] in the form of the SQUAD technique was extended in other works of the literature. The Continuous Zoom technique [17] provided better selection times when compared to Discrete Zoom [17], SQUAD and ray-casting. The Zoom and Expand techniques [18] also perform higher than the standard ray-casting and SQUAD techniques, but focus on dense environments. The Expand technique also provided better selection times than the Zoom. The similarities between these techniques are represented in our taxonomy. The reach of these techniques is screen-space, as they work on a 2D large scale display with a 3D perceived space and 3D user input.

The Go-Go technique [21] is the earliest approach on arm-extension techniques. It provides arms-reach selection in a 3D environment with a 3D input space. Only single objection is available, using a direct approach.

In World in Miniature [23], users have access to all objects in the 3D environment. Users control the copy of the environment using 3D input and can select objects using a direct approach. Even though the whole environment is available, there is a limit to its size, so we consider this techniques' reach as scaled. Voodoo Dolls [24] gives users the possibly of selecting an object using a direct approach in the 3D

Techniques	Perceived Space		Input Space		Reach				Cardinality		Strategy	
	2D	3D	2D	3D	Screen-Space	Arm-length	Scaled	Infinite	Single	Multiple	Direct	Progressive Refinement
Vacuum [1]	•		•					•		•	•	
Go-Go [21]		•		•		•			•		•	
Stretch Go-Go [4]		•		•				•	•		•	
Spotlight [32]		•		•				•	•		•	
Aperture [33]		•		•				•	•		•	
Orientation [33]		•		•				•	•		•	
SQUAD [16]		•		•	•				•			•
Discrete Zoom [17]		•		•	•				•			•
Continuous Zoom [17]		•		•	•				•			•
Zoom [18]		•		•	•				•			•
Expand [18]		•		•	•				•			•
World in Miniature [23]		•		•		•			•		•	
Voodoo Dolls [24]		•		•			•		•		•	
Prism + Go-Go [22]		•		•			•		•		•	
Scaled HOMER [31]		•		•			•		•		•	
3D Magic Wand [29]		•		•		•				•	•	
Serial [28]		•		•			•			•	•	
Parallel [28]		•		•			•			•	•	

Table 2.1: Classification of selection techniques.

environment to create a doll representing that object. It can select objects at distances greater than the reach of the users arms, using a pinching gesture. Both Prism + Go-Go [22] and the Scaled HOMER [31] improve previous work by increasing precision when interacting with objects. Both techniques only allow single object selection using a direct approach, but the reach of the techniques is considered scaled.

The Vacuum technique [1] stands out as the only technique with 2D input and that is used in a 2D perceived space and which allows infinite object selection using a 2D cone in a tabletop. Also gives the possibility of selection multiple object with a direct approach. The Stretch Go-Go [4] technique improves on the previously mentioned Go-Go, by being able to extend the virtual arm until infinite. The Spotlight [32] technique was developed as means to select objects without using a mathematical ray, using a cone instead. A bat is used as 3D input and a desktop is used to display the 3D environment. Even though more than one object can fall within the selection volume, only a single object can be selected, using a direct approach. The Aperture and Orientation techniques [33] were developed to advance the state of the art by using and HMD as their display. They improve the disambiguation method used in the Spotlight technique, by using either rotation of the object, and also they change from where the cone volume originates. Both technique can only select one object using a direct method, using the bat as the 3D input.

Even though we do not focus on multiple object selection, the following techniques provided good insight on what type of environments multiple selection techniques should be used. The 3D Magic Wand [29] technique was compared against the brush and lasso techniques for multiple object selection in IVEs. It gives users arms-reach selection, using a direct approach. Two other types of technique were developed for MOS, serial and parallel [28]. The serial techniques are only capable of selecting a single per operation, and on the parallel techniques more than object could be selected. One of the serial techniques used was ray-casting, making us consider this techniques' reach as infinite. All three techniques select the objects based on a direct approach.

Despite the various studies that have been done on object selection, there is still a shortcoming of approaches for out-of-reach object selection. Even though some techniques are capable of this, there is a lack of approaches that focus specifically on distant objects. That being said, with the proposed work of this dissertation we aim to tackle this challenge.

Chapter 3

Proposed Technique: PRECIOUS

In this chapter we will detail the implementation of our technique, PRECIOUS (Progressive Refinement using Conecasting in Immersive environments for Out-of-reach Selection). PRECIOUS is a combination of cone casting and progressive refinement. This technique reaches the far plane, allows multiple object selection and it is the first implementation of progressive refinement in virtual reality.

3.1 Technique Overview

PRECIOUS combines cone casting and progressive refinement to provide a quick and accurate selection technique. The cone volume is used to select the objects in the scene. By pointing in the direction of an object and closing his hand, the user can then open his hand to select the objects that fall inside the cone. The size and aperture of the cone can also be changed. We incorporate progressive refinement in our technique when the user makes a selection and three or more objects are inside the cone. In Section 5.1 we studied three travel techniques in VR and from the results of this test, we chose to use the Teleport technique [48] (also called infinite velocity).

Figure 3.1: Object Bounding box highlighted when inside the selection cone.

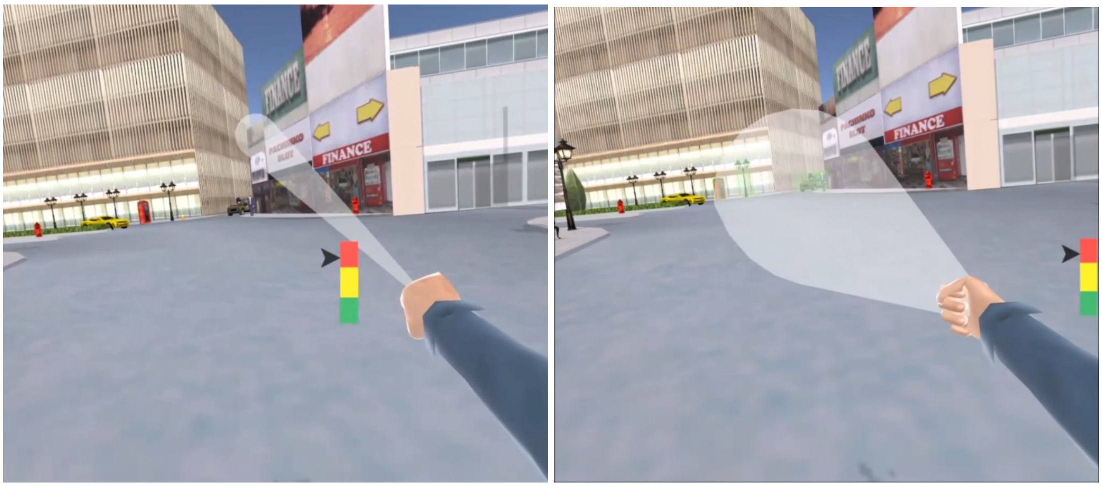


Figure 3.2: Using the wrist rotation to control the aperture of the cone.

To help users better understand which objects are inside the cone volume, their bounding boxes are shown as a partially visible green cube (Figure 3.1). While maintaining the hand closed, the user is able to modify the cone. To execute a selection action, the user has two choices: single or multiple object selection. To perform a single object selection the user simply opens his hand. Additionally, if a multiple section is to be performed, the user has to close his hand with more pressure. In the following sections we describe in greater detail how the selection process works on PRECIOUS.

3.2 Cone Manipulation

Cone manipulation is one of the key elements of our technique. There are two operations for manipulating the selection cone, one to control the aperture and another to control the reach of the cone. The Aperture Selection technique [33] employs a cone metaphor where the user can increase the size of the cone by moving his hand backwards and forwards. The vector used in this technique uses the direction from the eye point to the users hand to control the size of the cone. They use this vector since the origin of the cone coincides with the eye point of the user. We instead use the rotation of the users wrist (Figure 3.2), as our cones' origin is in the users dominant hand. When the wrist is rotated clockwise, the aperture of the cone increases until the opening angle reaches 15 degrees. If rotated in the opposite direction, the aperture will decrease until the opening angle is 7 degrees.

Also, similarly to Stretch Go-Go [4], and to help users identify in which region their hand current is, a widget is shown when the cone is active. The widget shows the three regions with an arrow pointing towards the one currently active (Figure 3.3 Left). We show this widget around the users hand instead of on the right side of the user. We made this distinction to the Stretch Go-Go as this approach usually does not work in VR as our eyes are unable to focus on something so close. Our approach uses a diegetic UI that follows the users right hand. And furthermore, when users are controlling the cone, the widget is always visible, as the hand is normally in front of the user. When users move their hand, they only need to give a quick glance to their hand to know what region they are in. Besides this improvement,

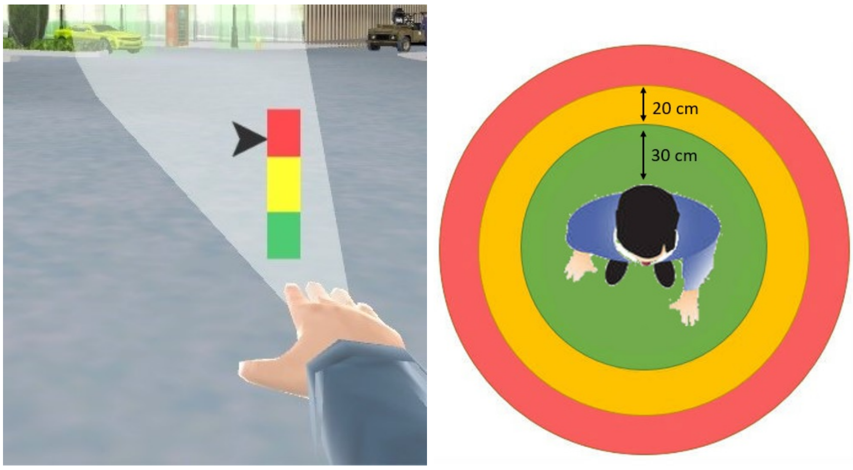


Figure 3.3: Left: Widget showing the regions. Right: Top view of the regions.

we also changed the vector used to control the position of the virtual hand. The direction vectors was directed from the right side of the torso, to allow users a more natural arm movement when withdrawing their hand to innermost region.

When manipulating the reach of the cone, we adopted a metaphor similar to Stretch Go-Go [4], with three regions around the users space, using the users natural hand position as the middle region (Figure 3.3 Right). When manipulating the reach of the cone, we adopted a metaphor similar to Stretch Go-Go [4], with three regions around the users space, using the right side of the users torso as the nearest position where the cone will retract. We considered this position more natural when using a cone as the selection volume, as users can still point the cone in the correct direction when decreasing the reach. When a user extends his hand into the outermost region, which is starts 50 cm from the user and corresponds to the red region, the cone stretches in that direction at 5 m/s. Placing the hand in the innermost region, depicted in green and which covers the first 30 cm around the user, causes the cone to decrease in size at the same 5 m/s. When operating the cone when the hand is placed in the middle region, represented in yellow and located between the two other regions, the size of the cone remains unchanged.

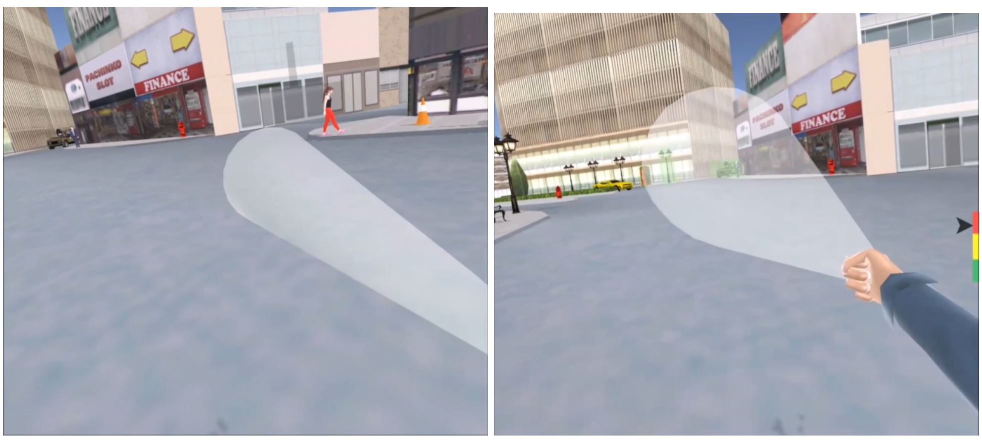


Figure 3.4: Controlling the reach of the cone.

Figure 3.5: Double Selection process.

3.3 Progressive Refinement

The next stage of our technique was based on the progressive refinement method [17].

When a user performs an object selection, one of three things can happen: the desired object is selected and the process ends; two objects are selected and they are shown directly in front of the user or the progressive refinement action is triggered. The first case is the terminal one, and is always the last step in the refinement process. On the second case, when two objects are inside the selection cone and the user performs a selection action those objects are shown in front of the user side by side (Figure 3.5). The rest of the objects are hidden during this step to make the selection of the object easier.

On the third case, when three or more objects are selected, we move the user close to the objects using the progressive refinement process. We cast a ray from the users hand position into the center of a sphere that contains all the objects currently inside the selection cone (Figure 3.6 Left). This sphere is calculated based on the bounding boxes of all objects that were inside the cone at the moment of selection. The intersection point of the ray with the sphere gives the position where the user will be moved to (represented in red in Figure 3.6 Center). To move the users we perform a simple Teleport action (Infinite Velocity), based on the results of a user evaluation described in Section 5.1. This ensures that users have all previously selected objects in their field of view when being moved to this new position (Figure 3.6 Right). The process will continue until only a single object is selected.

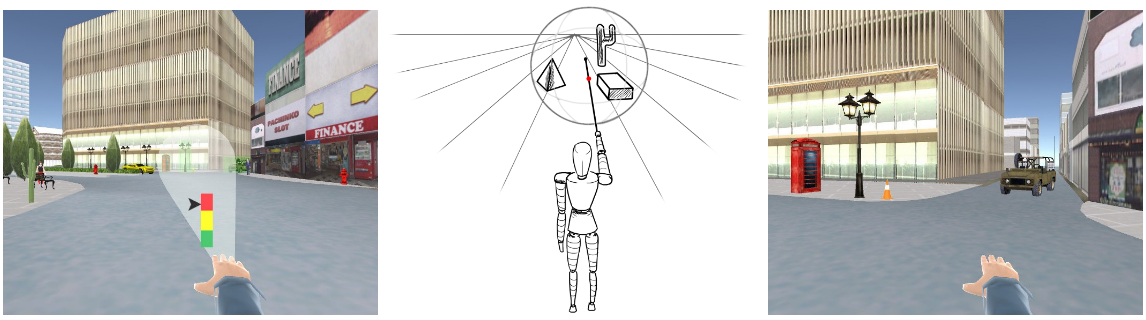


Figure 3.6: Refinement Process. Moving the user to his new position.

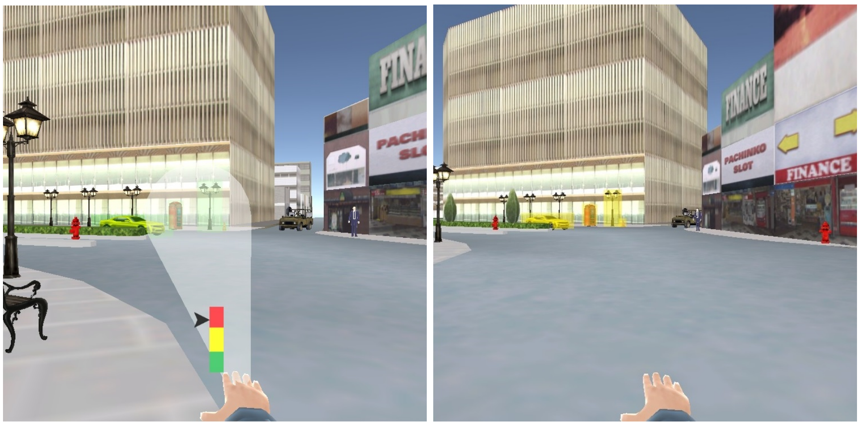


Figure 3.7: Example of multiple object selection.

3.4 Multiple Object Selection

Besides allowing single object selection, PRECIOUS also allows users to select multiple objects in just one operation. To trigger this action, users closes their hand with more pressure. All the objects that are inside the cone volume when a selection action is triggered will be selected (Figure 3.7), and will be highlighted with an orange box.

3.5 Usage Example

We will now provide an usage example of PRECIOUS, with the objective of selecting an out of reach object in our environment.

User start by aiming the selection cone in the direction of the object they intend to select (Figure 3.8 Left). If the object is too far away from the users current position, they can control the reach of the cone by moving their hand backwards and forwards. The three concentric regions around the user, represented as a *widget* in his hand. To reduce the reach of the cone, users put their hand in the innermost region. When users move their hand to the outermost region, the cone will increase in size, and lastly in the middle region the size of the cone remains unchanged (Figure 3.8 Center).

Figure 3.8: Left: Pointing to the desired objects. Center: Cone stretches and is now intersecting objects. Right: User was moved to the new position.

Figure 3.9: Left: Performing a double object selection. Center: Two selected objects are shown to the user. Right: Desired object selected.

The user also has control over the aperture of the cone, and can increase it by rotating his wrist in a clock-wise direction and decrease it in the opposite direction. When selecting an object in a semi-cluttered environment the user may select more than the object he intends. When this happens, we resort to the progressive refinement method, where the position of the user is changed to be closer to the selected objects (Figure 3.8 Right). If even after the refinement process the user still is not able to select the object, and can only select two object we make use of the double selection process (Figure 3.9 Left)). When this action is executed, the objects that were selected are shown in front of the user and the rest are hidden until a selection is made (Figure 3.9 Center). Finally, the user selects only the desired object and returns to his initial position with the object now selected and highlighted (Figure 3.9 Right).

3.6 Summary

The principal focus of the work of this dissertation consisted in the development of a technique for out-of-reach object selection in immersive virtual environments. In this chapter we presented an approach to this challenge, having used already existing techniques as a starting point.

In this approach the user controls a cone selection volume that can be manipulated to achieve a more accurate selection. Users can regulate the aperture of the cone using the rotations of the wrist. The reach of the cone can also be adjusted by moving the users dominant hand forwards or backwards from the right side of his torso.

When three or more objects are inside the cone we use the progressive refinement method to move the user closer to these objects. If two objects are inside the cone, they are shown to the user and the rest are hidden until a selection has been made. This approach lets users refine their selection until they select the desired object.

Chapter 4

Prototype

To adequately evaluate our proposed solution, a prototype was developed. In this section, we describe the architecture and setup used in this prototype. We will detail the important aspects of the architecture and the essential elements used in our setup.

4.1 Architecture

The development of our prototype followed the architecture represented in Figure 4.1. It consists of three listeners whose function is to collect data from the hardware and deliver it into our modules. The Hand Rotation Listener is responsible for capturing the data from our custom hardware for hand rotation

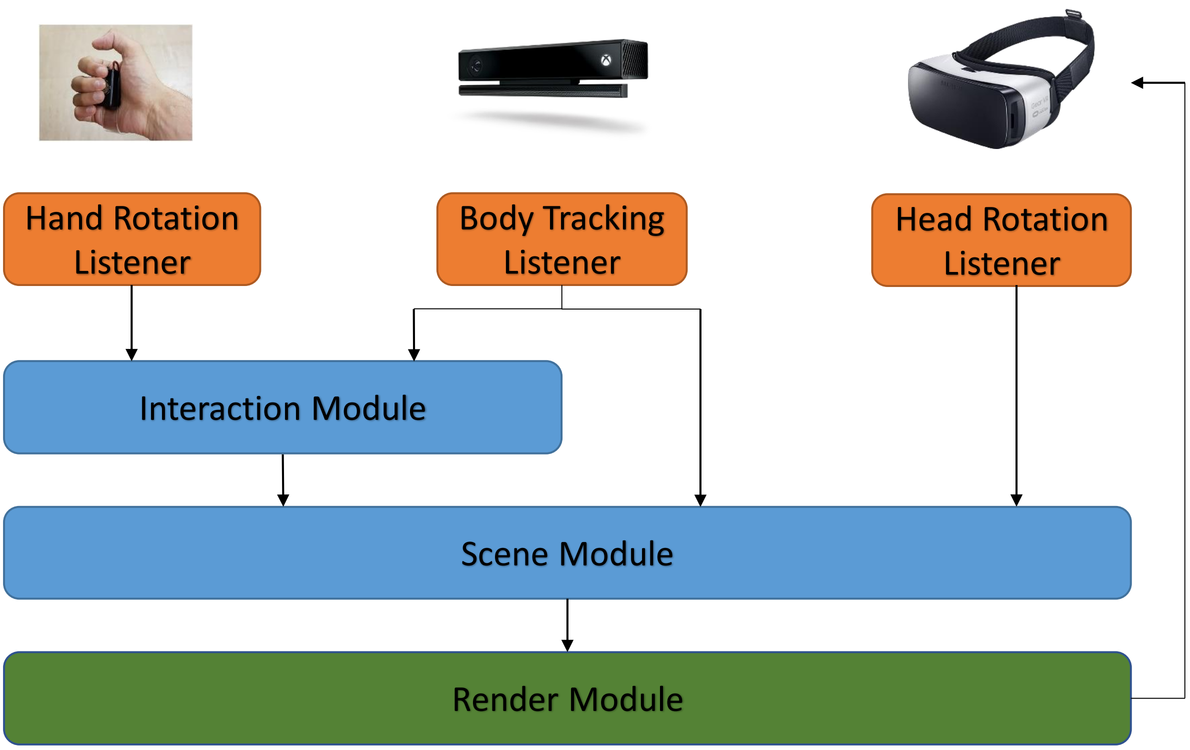


Figure 4.1: PRECIOUS Architecture.

tracking. The Body Tracking Listener gives us information regarding the position of the user. The Head Rotation Listener is in charge of sending the rotation of the users head to the scene module through the use of Kinect depth cameras. It also featured three modules that handle the raw data sent from the listeners. The Interaction Module is our main module, and handles data from the Hand Rotation Listener and Body Tracking Listener. It combines this data and correctly updates the position and rotation of the users' body. It is also responsible for handling the code from our technique. This data is then transmitted to the Render Module. The Scene Module receives data from the Body Tracking Listener and the Head Rotation Listener. This data is sent out to the Render Module which combined with the data from the Scene Module, is then sent to the Gear VR.

4.2 Setup

The setup used in our user evaluation is comprised of several non-evasive and affordable tracking hardware. We are able to capture the correct orientation of the user, as well has providing accurate hand and head rotation.

Hand Rotation Tracking

For hand rotation tracking we use an IMUduino, a custom hardware that includes an IMU and Bluetooth LE modules. The device is placed in the users hand using an acrylic clip (Figure 4.2). It also gives user control over 3 DOF for hand rotation and features a pressure pad, which can detect if the hand is open or closed. This pressure pad is also able to distinguish two pressure levels when being pressed. These pressure levels allowed us to separate what operation the user wanted to perform. The device communicates with a standalone script that sends the data over Wi-Fi to our head-mounted display.

Figure 4.2: IMUduino. Our custom made device for hand tracking.

Figure 4.3: Kinect V2 Camera.

Body Position Tracking

For body position tracking, we use three Kinect V2 depth cameras (Figure 4.3). Each Kinect camera is connected to a computer which analyses the data regarding 12 joints of the users skeleton. We also apply a double exponential smoothing filter[1], to try and hinder the jitter and noise from the data received. This data is then forwarded to our centralised unit, where the joint positions are converted into joint rotations, and then delivered to the Samsung Gear VR head-mounted display using Wi-Fi.

Head Tracking

For head tracking we use a Samsung GearVR (Figure 4.4) head-mounted display. On a first evaluation a Samsung Galaxy S7 smartphone was used, but given the problems that arose when using the IMUduino in conjunction with the tracking data from the Kinects', we chose to use a Samsung Galaxy S6 smartphone. It gives the user 96 degrees of field of view and is composed by a smartphone which has a resolution of 2560x1440 pixels and a refresh rate of 60 Hz. Thus, featuring motion sensors in a form of a gyroscope and an accelerometer to allow its user to move the head with 3 degrees of freedom inside the VE.

The Samsung Gear VR also served as a calibration instrument. The user was instructed to raise his left hand above his head and touch the pressure button located on the right side of the device with his right hand. When this action was performed, our Body position tracking unit would know which person in the room was the user.

Figure 4.4: Samsung Gear VR

[1] Skeletal Joint Smoothing White Paper: `http://msdn.microsoft.com/en-us/library/jj131429.aspx`, last visited 17th October 2016.

Tracking Consolidation

The data from all these sensors was integrated to correctly update the avatars position and also to accurately show the users head rotation in the virtual environment. The Body tracking unit communicated with the Body Tracking Listener, to correctly update the position and rotation of the avatar. It is also responsible for providing this data to the Interaction and Scene modules of our architecture, which is indispensable when changing the position of the users dynamically during the experiments. The Head tracking unit is responsible for accurately adjusting the rotation of the users head in the virtual environment. It is our head-mounted display, which is able to rendering the final image shown to the users. Hand rotation tracking was in control of handling both the rotations of the hand, and also to detect what type of touch the users did, using the pressure pad. The state of the users virtual hand was also changed with the data provided by this unit, as any change in state (open or closed hand) was represented in the virtual hand.

Our prototype was developed using the Unity3D engine with gravity and object collision disabled. The environment used in the evaluations consisted of a model of the city of Osaka, Japan, which was populated with objects that the users could select. And as previously mentioned, we also helped users better understand which objects they are interacting with by having their bounding box be shown as a partially visible green parallelepiped.

4.3 Summary

In this section we detail the important aspects of the architecture used during the development of this work, and its principal components. We also describe the setup used in our user evaluations, comprised of several hardware elements. We also explain how the integration of these sensors was done to provide accurate body, hand and head tracking to our application.

Chapter 5

Evaluation

During the course of this work, there were two evaluation moments, carried out with users and documenting their preferences and commentaries.

The first evaluation moment was carried out to determine the best way to move a user in a virtual environment. With the results from these tests, we developed a prototype for object selection, followed by a second evaluation moment. In this final user evaluation, we compared our proposed solution to two other approaches to object selection, with the objective of determining if combining progressive refinement and cone casting had significant advantages over the other approaches.

5.1 Travel Techniques Evaluation

We studied three different techniques for travel in immersive virtual environments. Our objective with this evaluation was to understand what impact these techniques had on the users, and which one was more suitable when changing the position of a user in a virtual environment.

Teleport Technique

The Teleport technique [48], also known as infinite velocity, translates a person instantaneously from their current position to the next checkpoint, as depicted in Figure 5.1A. Their results show that people were unable to process the information regarding the target direction accurately when using this technique.

Linear Motion

This technique consists of moving the user along a linear path for two seconds with a constant velocity, until the next checkpoint (Figure 5.1B). The velocity choice is based on previous work [44] and varies between 30 m/s and 50 m/s depending on the checkpoint distance.

Animated Teleport Box

We developed the Animated Teleport Box technique with the objective to combat the negative effects of the Teleport technique. Two 1.5 second animations were played when a user was being translated from their current position to next checkpoint. The first one animated the Box to rise up and surround the user, and the second one executed the same animation but in the inverse direction. The box has 2.3 meters on each side so that users would not feel too confined to the box when travelling. We employ the elevator metaphor, with the intention of not showing the users that they were being moved and as a means of decreasing the disorientation that might be felt after being teleported, as depicted in Figure 5.1C.

5.1.1 Task Design

To validate the techniques described above, we completed a user evaluation. Our aim was to understand which of the techniques were preferred by the users, and also the amount of impact that each had on users. The virtual environment was a model of the city of Osaka, Japan (Figure 5.2), which was populated with six coloured spherical checkpoints to where the users would be travelling to.

The path taken during the tests is depicted in Figure 5.3, where all six checkpoints and the users starting positions are represented. The distance between each of the checkpoints varied between 70 meters and 110 meters, making the total distance covered in each task about 500 meters.

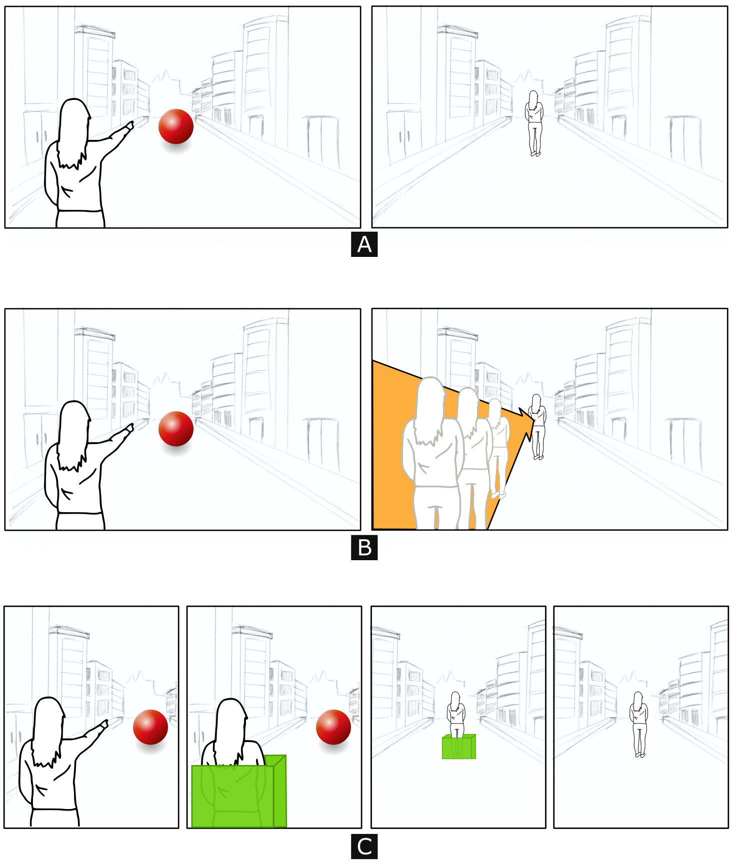


Figure 5.1: Travel techniques: A) Teleport; B) Linear Motion; C) Animated Teleport Box.

Figure 5.2: Virtual Environment.

5.1.2 Methodology

Each user evaluation session adopted the same protocol, starting the initial briefing with a quick explanation to the experiment and also with a description of the techniques. To avoid biased results from users becoming familiarised with the techniques and used to the environment, the techniques were presented in a partial random order, so all permutations were exhausted.

During each travel, the users were told where the next checkpoint would be (to their left or right) and were also instructed to point to said checkpoint before travelling using of the techniques. We gave this instruction to ascertain that if when users lost their frame of reference to the checkpoint they were pointing to. The users had no control over the path that they would take, and would only be in charge of pointing to the checkpoints. We allowed the users a one minute adjustment period to the environment of one one minute, before travelling to the first checkpoint, to make sure they knew where they were and where they were being moved to. Each session took on average thirty minutes, which ended with a brief questionnaire about their experience.

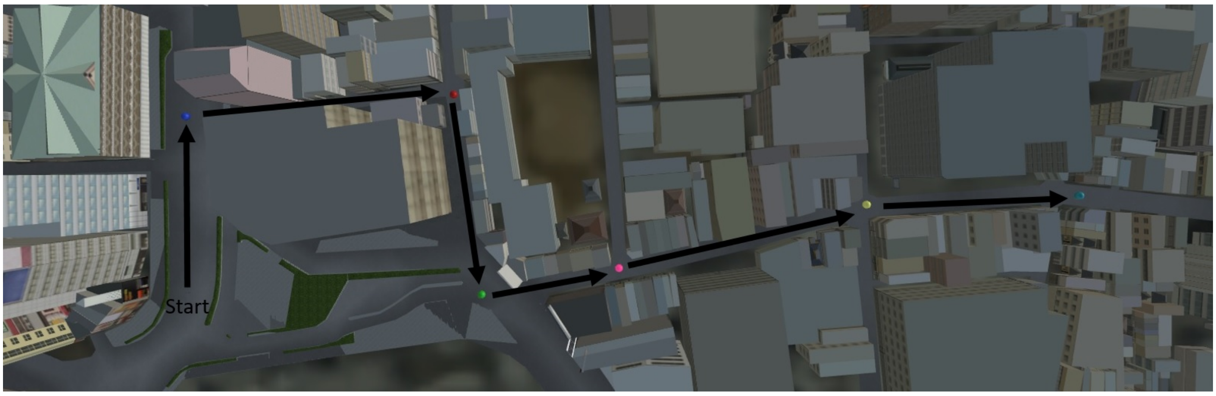


Figure 5.3: Top-view of the Virtual Environment with the path travelled.

5.1.3 Apparatus and Participants

We tested the techniques in our laboratory in a controlled environment. In our experiment we used a Samsung GearVR head-mounted display. The Samsung GearVR has 96 degrees of field of view and is composed by a Samsung Galaxy S7 smartphone which has a resolution of 2560x1440 pixels and a refresh rate of 60 Hz. Thus, featuring motion sensors in a form of a gyroscope and a accelerometer to allow its user to move the head with 3 degrees of freedom inside the VE.

Twenty participants completed the user evaluation, two of them were females (Figure 5.4). The participants' age ranged from 19 to 31 years old, with a mean of 24. Seven of the participants already had previous experience with virtual reality visualisation.

5.1.4 Results and Discussion

We present the main observations made during the evaluation sessions, as well as the difficulties and suggestions communicated by the users. Additionally we discuss the analysis and the results obtained.

We first conducted a Shapiro-Wilk to test our data for normality. Afterwards a Friedman non-parametric test was conducted to look for statistical significance between the three tested techniques. When statistical differences were found, we conducted a Wilcoxon Signed-Ranks Test to look for statistical significance on each pair of techniques with the Bonferroni correction. For a better comparison regarding task performance, we subtracted the animation times from the total time following the formula: $T' = T - \alpha \times (n-1)$, where T is the total time, α the path time (3 seconds in the animated box, 2 on the linear motion, and zero on the teleport) and n the number of travels (6 in our case). This represents the recovery time that users needed after each travel. Figure 5.5 shows both the total time of

Figure 5.4: Participant during a user test.

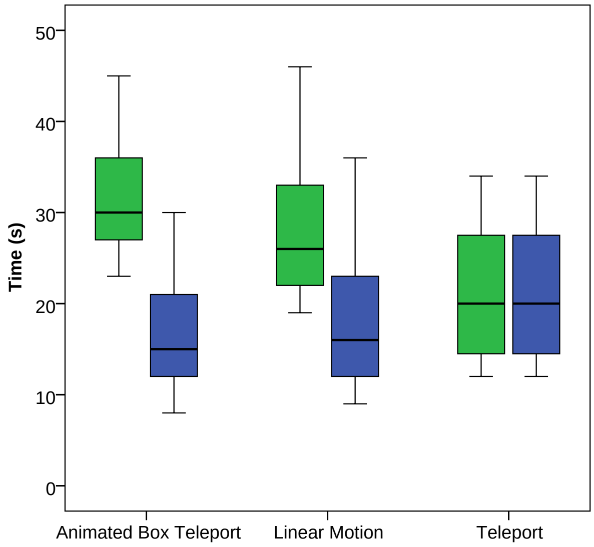


Figure 5.5: Box-plot representing the time elapsed on each task. Green box-plots represent total time, and blue the time excluding techniques' animations.

each travel and the recovery time needed by users after each travel. By looking at Figure 5.5 we can notice a slightly better performance on the Animated Teleportation Box technique, but without statistical significance. Because of that, we can state that efficiency is similar in all the tested techniques.

Regarding questionnaires' data (Table 5.1) we found that users felt more physical discomfort using the Linear Motion Technique (Z=-2.699, $p < 0.01$ against Animated Teleport Box and Z=-2.386, p=0.017 against the Teleport). Despite of the discomfort caused by the Linear Motion, participants stated it as their favourite technique in most cases, Because of the similarity between user preferences on both Animated Box Teleport and Teleport we conducted an additional test on the total times of the test task. This test confirms a better result on such condition with the Teleport as it does not need additional time among the movement between positions (Z=-3.114, $p < 0.01$ between Animated Box and Z=-2.578, p=0.01 against Linear Motion).

5.2 PRECIOUS Validation

To validate our technique we compared it against two techniques from the literature, Stretch Go-Go [21] and Spotlight [32]. We implemented these technique in our prototype as they provide a good baseline for the tests,as they are adequate for out-of-reach selection and capable of infinite reach.

Question	**Animated Teleport Box**	**Linear Motion**	**Teleport**
It was easy	5 (1)	5 (1)	5 (1)
I was satisfied	4 (2)	4.5 (2)	4 (2)
I felt physical discomfort*	1 (1)	2 (3)	1 (1)
I felt visual discomfort	1 (1)	2 (2)	1 (1)

Table 5.1: User preferences: Median (Interquartile Range). * indicates statistical significance.

Figure 5.6: Stretch Go-Go technique.

5.2.1 Stretch Go-Go

The Stretch Go-Go technique was developed as a means to overcome the limitations of the Go-Go technique [21]. In the latter, the reach of the virtual hand in the virtual environment was limited by the length of the user's arm. The ”stretch go-go” technique allows infinite stretching of the virtual hand using only arm motion for control, so that all objects in the environments can be reached. The space around the user is divided into three concentric regions. The user's natural hand position is considered to be in the middle region. As the users stretch their hand into the outermost region, the arm begins to grow at a constant speed in that direction. When the arm is moved to the innermost region, the arm retreats at that speed. In the middle region, the arm length remains the same. By mapping the position of the physical hand to the virtual hand velocity, any arm length can be achieved.

To aid the users, a gauge is shown on their dominant hand that indicates in what region they are currently operating (Figure 5.6). We added an additional exponential smoothing filter to this technique, to increased the control of the position of the virtual hand,. This position is calculated using the following formula: $Vpos' = 0.5 * Vpos - 0.5 * (handPos + vHandDirection) * vDist)$, where Vpos represents the current position of the virtual hand in the environment, handPos is the position of the users hand in the virtual environment, vHandDirection is the direction of the vector from the users origin to the users hand and vDist represents the distance from the users origin where the virtual hand is.

5.2.2 Spotlight

The Spotlight technique was intended to overcome the main issue of ray-casting, which is low accuracy when targeting objects that are out-of-reach from the user [32]. The technique employs a method called ”spotlight selection” in order to make the selection softer than a mathematical ray, so that some error was tolerated when aiming. A bat was used as a prop when using the selection volume. A selection cone is first established with its apex at the bat and its axis at the X-axis of the bat. Objects that fall inside the

Figure 5.7: Spotlight technique.

cone are candidates for selection. When more than one object is inside the cone an anisotropic distance metric is calculated as a disambiguation mechanism. If two objects are at equal Euclidean distance from the bat, the closest to the center line of the cone is selected. If otherwise the two objects produce the same angle at the bat with the center line, the object with the smaller Euclidean distance is selected. Figure 5.7 shows the visual feedback given by the Spotlight technique. Users can see which objects are candidates for selection as their bounding box be shown as a partially visible green parallelepiped.

5.2.3 Methodology

All user evaluation followed the same format and usually lasted 45 minutes. The experiment started with an explanation of the techniques that would be tested. To help users better understand the environment and how each technique worked, a small video was also played during the explanation.

Afterwards, participants were given three minutes before each task as a training period to adjust themselves to the environment and to the technique being tested. The order of the techniques being tested was presented in a partial random order, so that all permutations were exhausted and to avoid biased results as users could become familiarised with the approach and the environment. Each participant was instructed to perform four tasks, described in the next section.

After completing the tasks for each technique, participants were asked to fill out a questionnaire about the approach tested. To finish the experiment, participants were asked to fill out a profiling questionnaire.

5.2.4 Tasks

Participants were requested to complete four tasks for each technique. They all consisted of selecting a cactus in our virtual environment (Figure 5.8). With the exception of the buildings that composed the environment, all objects were selectable. In order to avoid long session times, we restricted the duration of each task to three minutes. If participants reach the time limit they would be informed they could stop, and we took this as an unsuccessful attempt.

(a) Task 1

(b) Task 2

(c) Task 3

(d) Task 4

Figure 5.8: Tasks performed by the participants.

The tasks were designed to be increasingly difficult, as the cactus would be placed further away from the users initial position or surrounded by more objects. Every time users selected an object that was not the cactus, we would log it as an incorrect selection.

5.2.5 Apparatus and Participants

The experiment was carried out in our laboratory in a controlled environment (Figure 5.9), using the setup that was specified in Section 4.2. In this evaluation we had a total of 18 participants (2 female), with ages between 18 and 40 years old with the majority (62%) being between 18 and 25. The majority held at least a Bachelor degree (62%). When asked if they had ever experience virtual reality, 39% said that they never had, and only 28% of users said that never used a gesture recognition system, such as the Microsoft Kinect or the Wii Remote.

5.2.6 Object Selection Evaluation

A user study was conducted to properly evaluate the quality of the approach developed against those from the literature.

Figure 5.9: Participant of the experiment.

During the experiment we gathered objective data, with the use of a logs of each task and subjective data in the form of questionnaires. The logs registered information regarding the completion time for each task and the number of incorrect selection made. Additionally, we also calculated the success rate for each user evaluation, and would consider it a failed test if the user exceeded the time limit. The first test we performed was the Shapiro-Wilk test, to ascertain the normality of the data. A repeated measures ANOVA test with a Greenhouse-Geisser correction was then carried out to find significant differences in normal distributed data. We used this test for three out of the four tasks. Additional Friedman non-parametric tests with Wilcoxon Signed-Ranks post-hoc tests were also conducted. In both the ANOVA and the Friedman non-parametric tests post-hoc tests used Bonferroni correction (corrected sig. = sig x 3).

5.2.7 Objective Data Results

We measured the total time that participants took on each task, taken in seconds, and depicted in Figure 5.10. We also registered the number of incorrect selections, described in Table 5.2. The success rate of the techniques was also analysed, and is detailed in Figure 5.11

We found statistical significance in the completion time of all tasks (Task 1: $\chi^2(2)$=17,375, p$<$.0005; Task 2: F(1.013,8.102)=18.327, p=.003; Task 3: $\chi^2(2)$=19, p$<$.0005; Task 4: t(9)=-3.802, p=.004). In regards to the number of incorrect selections, we found statistical significance only in Task 4: $\chi^2(2)$=25.064, p$<$.0005.

When comparing the completion times in the first task, post-hoc test showed that the Spotlight approach (avg=10s) was faster than both PRECIOUS (avg=18s, Z=-2.430, p=.045) and Stretch Go-Go (avg=63s, Z=-3.479, p=.003) and PRECIOUS to be faster than Stretch Go-Go (Z=-3.574, p$<$.0005).

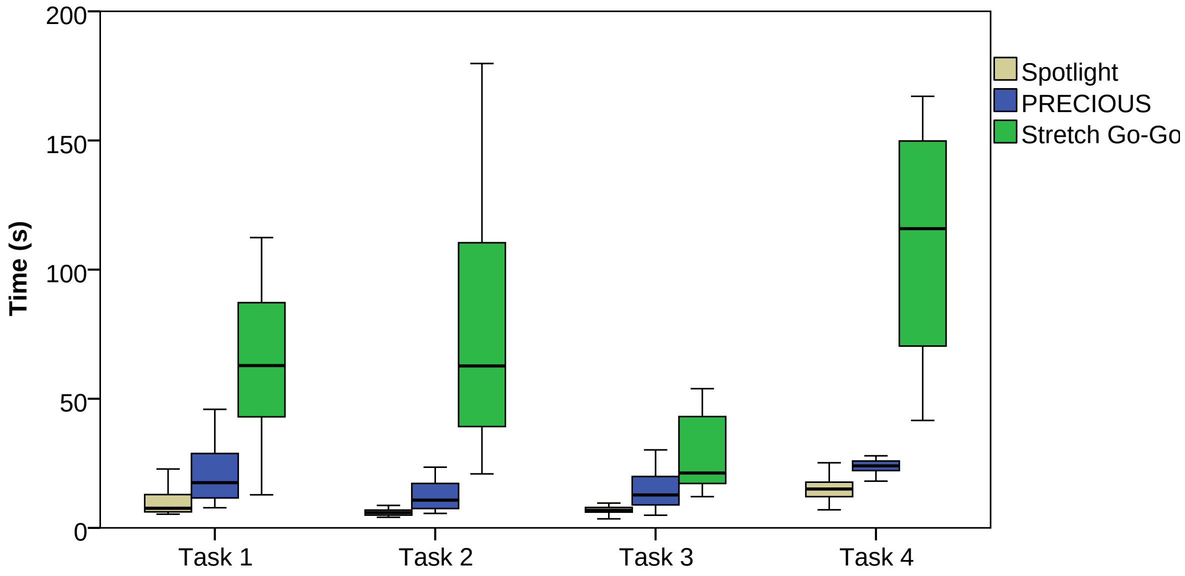


Figure 5.10: Completion time of the four tasks using the three techniques, in seconds. The graphic presents the median, first and third inter quartile ranges (boxes) and 95% confidence interval (whiskers).

For the second task, Spotlight was also faster than other two approaches, PRECIOUS (avg=11s, Z=-2.926, p=.009) and Stretch Go-Go (avg=63s, Z=-2.666, p=.024). PRECIOUS had significantly better completion times when compared to Stretch Go-Go (Z=-2.9340, p=0.009). This tasks reveals the flaws associated to the Stretch Go-Go technique, as the object was positioned further away from the user, the success rate dropped to 61.1%.

In the third task, the Spotlight (avg=7s) was faster than PRECIOUS (avg=11s, Z=-3.030, p=.006) and also Stretch Go-Go (avg=21s, Z=-3.296, p=.003). In this task, PRECIOUS also showed better results when compared to Stretch Go-Go (Z=-2.480, p=0.039). As expected, when the object is moved closer to the user the success rate of Stretch Go-Go increased to 83.3%.

On task number four, the Spotlight approach (15s) was also the faster than PRECIOUS (avg=24s, Z=-3.21, p=.003) and Stretch Go-Go (avg=115s, Z=-3.317, p=.003). As in the previous tasks, Stretch Go-Go proved to provide worse results than PRECIOUS (Z=-3.233, p=0.06). To analyse this task we had to conduct a Paired T-Test, as there were insufficient data from Stretch Go-Go for a Friedman test. This task also reveals that when objects are too close together some incorrect selections occur when using the Spotlight technique. In this task, the success rate of Stretch Go-Go is very inferior to the other techniques and there were only 4 participants that were able to complete this task before the time limit was met.

Technique	**Task 1**	**Task 2**	**Task 3**	**Task 4**
Stretch Go-Go	0 (0)	0 (0)	0 (0)	0 (0)
Spotlight	0 (0)	0 (0)	0 (0)	1.5 (3)
PRECIOUS	0 (0)	0 (0)	0 (0)	0 (1)

Table 5.2: Number of incorrect selections for each technique: Median (Inter quartile Range).

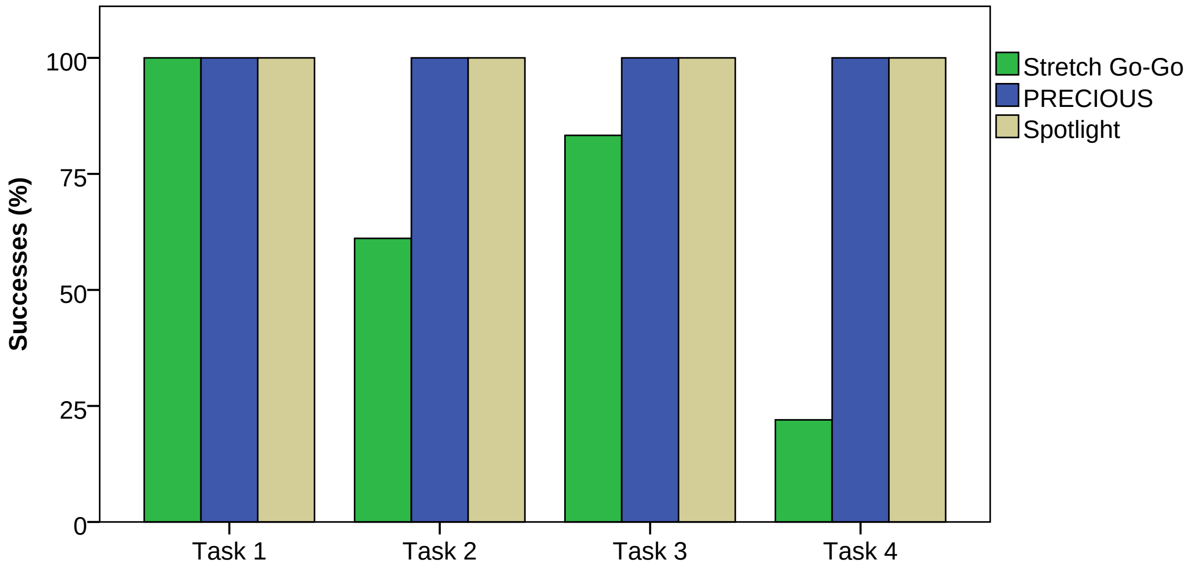


Figure 5.11: Success rate for the four tasks using the three techniques, in seconds. The graphic presents the median, first and third inter quartile ranges (boxes) and 95% confidence interval (whiskers).

5.2.8 Subjective Data Results

On the questionnaires users were asked about their experience with each technique. They were asked about the difficulty of the techniques, the fun factor, if they felt tired and if there was any discomfort. Additionally they were asked if the control of the cone in the Spotlight and PRECIOUS was difficult, and if controlling the virtual hand in Stretch Go-Go was easy or not. The questions were asked in the form of a Likert Scale from to 1 to 5, 5 being the favourable value, and the answers are depicted in Table 5.3. Both the Median and the Inter quartile Range are represented in the table.

When analysing the results from the questionnaires, we identified significant differences in ease of use (χ^2(2)=23.524, p<.0005), fun factor (χ^2(2)=27.180, p<.0005), fatigue felt (χ^2(2)=18.582, p<.0005) and the discomfort felt (χ^2(2)=22.189, p<.0005). Participants heavily recognise that Stretch Go-Go was harder to use (Spotlight: Z=-3,673, p<.0005, PRECIOUS: Z=-3,556, p<.0005), also less fun (Spotlight: Z=-3,660, p<.0005, PRECIOUS: Z=-3.572, p<.0005), more tiring (PRECIOUS: Z=-3.441, p=.003) and more discomforting (Spotlight: Z=-3.342, p=.003, PRECIOUS: Z=-3.475, p=.003).

In regards to the questions about the difficulty of using the cone in the Spotlight technique, participants responded positively ($\bar{x} = 4$, iqr=1). When questioned about the easiness of controlling the virtual hand, the answers corroborated the previous results ($\bar{x} = 1.5$, iqr=2). When answering the questions regarding PRECIOUS' cone, users did not give important significance to the control of the aperture ($\bar{x} = 3$, iqr=1). Controlling the reach of the cone was answered affirmatively by most users ($\bar{x} = 4$, iqr=2). The teleport technique used to move the users in each test was answered positively ($\bar{x} = 5$, iqr=0).

Question	Stretch Go-Go	Spotlight	PRECIOUS
It was easy *	1 (1)	4.5 (1)	4 (1)
I was satisfied *	2 (1)	5 (1)	4 (1)
I felt physical discomfort *	2.5 (2)	5 (1)	5 (1)
I felt visual discomfort *	3 (1)	5 (1)	5 (1)

Table 5.3: User preferences: Median (Inter quartile Range). * indicates statistical significance.

5.2.9 Discussion

It is observable that Stretch Go-Go is an ineffective approach when objects are out of the users reach. The time it takes to stretch the virtual hand plays an important part in the success rate of this technique, as when users missed the object they needed to adjust the position of the virtual hand. In the final task is where this is more evident, as the object is placed further away from the user where controlling the virtual hand becomes impractical.

Each task had an increasing level of difficulty, as the object would either be placed further away from the user (Task 2 and 4) or closer to other objects in environment, making it harder to perform a selection without errors (Task 3 and 4). Adding the data regarding the number of incorrect selection we can assert that even though the Spotlight approach is faster, it is more prone to errors when the difficulty of the task increases. Throughout all tasks Spotlight shows that it is the faster technique, even though some selection errors are present when the difficulty of the task increases. In some selections tasks, having an unwanted selection can have a great impact on the results.

All the objects in the environment had a box collider as their bounding box. Some comments by users were made in this regard, as some objects' bounding box were larger than their model, which could have had some impact in the tasks.

The complexity of PRECIOUS, mainly the need to increase and decrease the reach of the cone, had an impact in the completion time. We consider that the combination of progressive refinement with cone selection to be a superior technique, but the added difficulty of controlling the reach and the aperture of the cone pose a greater handicap than previously though. We can also attribute a higher completion time since there is an added time to travel between selections, due to the progressive refinement process.

5.3 Summary

In this chapter two evaluations of the work produced were described. The first focused on the evaluation of travel technique in immersive virtual environments. The Teleport technique was the one that provided the best results, both in terms of speed and impact of cybersickness on users.

This technique was incorporated in our selection technique for out-of-reach objects, PRECIOUS. To correctly determine if our approach to object selection of distant objects was superior to current approaches, a second evaluation moment was conducted. PRECIOUS was compared to the Spotlight and Stretch Go-Go techniques, to ascertain which would wield better results on different selection scenarios, both in terms of selection time and number of incorrect selections. Even though the Spotlight technique provided lower selection times across all tasks, some selections errors occurred when the selection task was more difficult. The Stretch Go-Go technique showed consistently higher selection times across all tasks, and proved inadequate to use when objects were very distant.

Chapter 6

Conclusions

The task of selecting an object is present in our everyday life. From desktop interfaces to virtual environments, there is a need to select an object before changing it in any way. With the recent research in the new, more affordable hardware, there is a greater interest in exploring approaches for object interaction in virtual environments.

When considering the current approaches for object selection in immersive virtual environments, few tackle the problem of selecting objects at great distances. The most common approaches of ray-casting and arm-extension suffer from jitter problems when the intended object is too far from the position of the user. Volume selection techniques on the other hand can more effectively deal with this problem, but when the environment is cluttered some unwanted selections may occur.

To deal with these problems, we developed PRECIOUS, a combination of progressive refinement and cone casting that allows users to select objects at various distances. To incorporate progressive refinement in our selection process, by allowing the user to select more than one object using the cone volume, and then refine his selection with those fewer objects. We moved the user in the virtual environment using Teleport, a travel technique for virtual reality [48]. This technique was chosen based on an evaluation conducted, with the objective of determining the optimal way to move a user in a virtual environment.

A formal user evaluation was then conducted, where PRECIOUS was compared with two other selection techniques. With the results from this evaluation we determined that Stretch Go-Go is an impractical technique when selecting objects that are very distant from the user (further than the size of a room). On the other hand we leaned that the Spotlight technique can provide faster completion times on standard selection tasks, but when the environment is not sparse it is prone to incorrect selections. Regarding PRECIOUS we can state that, even though it was not the fastest of the three techniques, the lack of errors and the uniform completion times across all scenarios tested, make it a suitable technique for object selection.

6.1 Future Work

In this work there were some aspects that can be considered worthy of improvements or alternatively can originate future work:

- **Adjust the control of the selection cone**

  From the results of the user evaluation, a possible next stage could be adjusting the type of control on the selection cone. Either removing the control over the reach of the cone, and having its reach be infinite, or by fine tuning the aperture of the cone.

- **Improvement to Stretch Go-Go**

  The Stretch Go-Go technique is not efficient when the selections involved distant objects. Some improvements can be made to overcome this, such as changing the virtual hand and adopting a spherical volume that scales the further it is from the user. Changing the direction vector used and the approach used to bring the hand closer and further away from the user can also be done to arrive at a faster and more reliable technique.

- **Investigate other progressive refinement metaphors**

  We combined cone casting and progressive refinement for the first time in virtual reality. There are different progress refinement techniques that can be explored [17], such as the Expand technique.

- **Combine PRECIOUS with manipulation approaches**

  A natural next step is combining PRECIOUS with technique for object manipulation. This would create an all around approach for object manipulation at any distance.

- **Explore final refinement stage with more objects**

  The last phase of PRECIOUS' progressive refinement shows the user the two objects he selected in front of him. This value can be customisable. Increasing it to larger values, and showing them in a barrel arrangement as described in [49] can increase the selection speed.

Bibliography

[1] A. Bezerianos and R. Balakrishnan. The vacuum: facilitating the manipulation of distant objects. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 361–370. ACM, 2005.

[2] J. Rekimoto. Pick-and-drop: a direct manipulation technique for multiple computer environments. In *Proceedings of the 10th annual ACM symposium on User interface software and technology*, pages 31–39. ACM, 1997.

[3] A. Cohé, F. Dècle, and M. Hachet. tbox: a 3d transformation widget designed for touch-screens. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3005–3008. ACM, 2011.

[4] D. A. Bowman and L. F. Hodges. An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In *Proceedings of the 1997 symposium on Interactive 3D graphics*, pages 35–ff. ACM, 1997.

[5] O. Hilliges, S. Izadi, A. D. Wilson, S. Hodges, A. Garcia-Mendoza, and A. Butz. Interactions in the air: adding further depth to interactive tabletops. In *Proceedings of the 22nd annual ACM symposium on User interface software and technology*, pages 139–148. ACM, 2009.

[6] S. Izadi, S. Hodges, S. Taylor, D. Rosenfeld, N. Villar, A. Butler, and J. Westhues. Going beyond the display: a surface technology with an electronically switchable diffuser. In *Proceedings of the 21st annual ACM symposium on User interface software and technology*, pages 269–278. ACM, 2008.

[7] A. Banerjee, J. Burstyn, A. Girouard, and R. Vertegaal. Pointable: an in-air pointing technique to manipulate out-of-reach targets on tabletops. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, pages 11–20. ACM, 2011.

[8] G. De Haan, M. Koutek, and F. H. Post. Intenselect: Using dynamic object rating for assisting 3d object selection. In *IPT/EGVE*, pages 201–209. Citeseer, 2005.

[9] R. J. Teather and W. Stuerzlinger. Pointing at 3d targets in a stereo head-tracked virtual environment. In *3D User Interfaces (3DUI), 2011 IEEE Symposium on*, pages 87–94. IEEE, 2011.

[10] D. Mendes, F. Fonseca, B. Araujo, A. Ferreira, and J. Jorge. Mid-air interactions above stereoscopic interactive tables. In *3D User Interfaces (3DUI), 2014 IEEE Symposium on*, pages 3–10. IEEE, 2014.

[11] P. Song, W. B. Goh, W. Hutama, C.-W. Fu, and X. Liu. A handle bar metaphor for virtual object manipulation with mid-air interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1297–1306. ACM, 2012.

[12] T. Grossman and R. Balakrishnan. The design and evaluation of selection techniques for 3d volumetric displays. In *Proceedings of the 19th annual ACM symposium on User interface software and technology*, pages 3–12. ACM, 2006.

[13] D. Vogel and R. Balakrishnan. Distant freehand pointing and clicking on very large, high resolution displays. In *Proceedings of the 18th annual ACM symposium on User interface software and technology*, pages 33–42. ACM, 2005.

[14] J. L. Reisman, P. L. Davidson, and J. Y. Han. A screen-space formulation for 2d and 3d direct manipulation. In *Proceedings of the 22nd annual ACM symposium on User interface software and technology*, pages 69–78. ACM, 2009.

[15] F. Argelaguet and C. Andujar. Visual feedback techniques for virtual pointing on stereoscopic displays. In *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology*, pages 163–170. ACM, 2009.

[16] R. Kopper, F. Bacim, D. Bowman, et al. Rapid and accurate 3d selection by progressive refinement. In *3D User Interfaces (3DUI), 2011 IEEE Symposium on*, pages 67–74. IEEE, 2011.

[17] F. Bacim, R. Kopper, and D. A. Bowman. Design and evaluation of 3d selection techniques based on progressive refinement. *International Journal of Human-Computer Studies*, 71(7):785–802, 2013.

[18] J. Cashion, C. Wingrave, and J. J. LaViola Jr. Dense and dynamic 3d selection for game-based virtual environments. *IEEE transactions on visualization and computer graphics*, 18(4):634–642, 2012.

[19] M. Veit, A. Capobianco, and D. Bechmann. Influence of degrees of freedom's manipulation on performances during orientation tasks in virtual reality environments. In *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology*, pages 51–58. ACM, 2009.

[20] R. Wang, S. Paris, and J. Popović. 6d hands: markerless hand-tracking for computer aided design. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pages 549–558. ACM, 2011.

[21] I. Poupyrev, M. Billinghurst, S. Weghorst, and T. Ichikawa. The go-go interaction technique: non-linear mapping for direct manipulation in vr. In *Proceedings of the 9th annual ACM symposium on User interface software and technology*, pages 79–80. ACM, 1996.

[22] C. Auteri, M. Guerra, and S. Frees. Increasing precision for extended reach 3d manipulation. *International Journal of Virtual Reality*, 12(1):66–73, 2013.

[23] R. Stoakley, M. J. Conway, and R. Pausch. Virtual reality on a wim: interactive worlds in miniature. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 265–272. ACM Press/Addison-Wesley Publishing Co., 1995.

[24] J. S. Pierce, B. C. Stearns, and R. Pausch. Voodoo dolls: seamless interaction at multiple scales in virtual environments. In *Proceedings of the 1999 symposium on Interactive 3D graphics*, pages 141–145. ACM, 1999.

[25] J. S. Pierce, A. S. Forsberg, M. J. Conway, S. Hong, R. C. Zeleznik, and M. R. Mine. Image plane interaction techniques in 3d immersive environments. In *Proceedings of the 1997 symposium on Interactive 3D graphics*, pages 39–ff. ACM, 1997.

[26] H. Benko and A. D. Wilson. Multi-point interactions with immersive omnidirectional visualizations in a dome. In *ACM International Conference on Interactive Tabletops and Surfaces*, pages 19–28. ACM, 2010.

[27] M. R. Mine, F. P. Brooks Jr, and C. H. Sequin. Moving objects in space: exploiting proprioception in virtual-environment interaction. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 19–26. ACM Press/Addison-Wesley Publishing Co., 1997.

[28] J. F. Lucas. Design and evaluation of 3d multiple object selection techniques. 2005.

[29] R. Stenholt. Efficient selection of multiple objects on a large scale. In *Proceedings of the 18th ACM symposium on Virtual reality software and technology*, pages 105–112. ACM, 2012.

[30] G. J. Wills. Selection: 524,288 ways to say "this is interesting". In *Information Visualization'96, Proceedings IEEE Symposium on*, pages 54–60. IEEE, 1996.

[31] C. Wilkes and D. A. Bowman. Advantages of velocity-based scaling for distant 3d manipulation. In *Proceedings of the 2008 ACM symposium on Virtual reality software and technology*, pages 23–29. ACM, 2008.

[32] J. Liang and M. Green. Jdcad: A highly interactive 3d modeling system. *Computers & graphics*, 18 (4):499–506, 1994.

[33] A. Forsberg, K. Herndon, and R. Zeleznik. Aperture based selection for immersive virtual environments. In *Proceedings of the 9th annual ACM symposium on User interface software and technology*, pages 95–96. ACM, 1996.

[34] D. A. Bowman, D. Koller, and L. F. Hodges. A methodology for the evaluation of travel techniques for immersive virtual environments. *Virtual reality*, 3(2):120–131, 1998.

[35] D. A. Bowman, E. Kruijff, J. J. LaViola Jr, and I. Poupyrev. *3D user interfaces: theory and practice*. Addison-Wesley, 2004.

[36] E. A. Suma, S. L. Finkelstein, M. Reid, S. V. Babu, A. C. Ulinski, and L. F. Hodges. Evaluation of the cognitive effects of travel technique in complex real and virtual environments. *Visualization and Computer Graphics, IEEE Transactions on*, 16(4):690–702, 2010.

[37] D. A. Bowman, E. Kruijff, J. J. LaViola Jr, and I. Poupyrev. An introduction to 3-d user interface design. *Presence: Teleoperators and virtual environments*, 10(1):96–108, 2001.

[38] S. P. Smith and T. Marsh. Evaluating design guidelines for reducing user disorientation in a desktop virtual environment. *Virtual Reality*, 8(1):55–62, 2004.

[39] D. A. Bowman, E. T. Davis, L. F. Hodges, and A. N. Badre. Maintaining spatial orientation during travel in an immersive virtual environment. *Presence: Teleoperators and Virtual Environments*, 8 (6):618–631, 1999.

[40] E. D. Ragan, A. Wood, R. P. McMahan, and D. A. Bowman. Trade-offs related to travel techniques and level of display fidelity in virtual data-analysis environments. In *ICAT/EGVE/EuroVR*, pages 81–84, 2012.

[41] S. Davis, K. Nesbitt, and E. Nalivaiko. Comparing the onset of cybersickness using the oculus rift and two virtual roller coasters. In *Proceedings of the 11th Australasian Conference on Interactive Entertainment (IE 2015)*, volume 27, page 30, 2015.

[42] J. J. LaViola Jr. A discussion of cybersickness in virtual environments. *ACM SIGCHI Bulletin*, 32 (1):47–56, 2000.

[43] C.-T. Lin, S.-W. Chuang, Y.-C. Chen, L.-W. Ko, S.-F. Liang, and T.-P. Jung. Eeg effects of motion sickness induced in a dynamic virtual reality environment. In *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 3872–3875. IEEE, 2007.

[44] R. H. So, W. Lo, and A. T. Ho. Effects of navigation speed on motion sickness caused by an immersive virtual environment. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 43(3):452–461, 2001.

[45] A. S. Fernandes and S. K. Feiner. Combating vr sickness through subtle dynamic field-of-view modification. In *2016 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 201–210. IEEE, 2016.

[46] D. A. Bowman and L. F. Hodges. Formalizing the design, evaluation, and application of interaction techniques for immersive virtual environments. *Journal of Visual Languages & Computing*, 10(1): 37–53, 1999.

[47] T. Grossman and D. Wigdor. Going deeper: a taxonomy of 3d on the tabletop. In *Horizontal Interactive Human-Computer Systems, 2007. TABLETOP'07. Second Annual IEEE International Workshop on*, pages 137–144. IEEE, 2007.

[48] D. A. Bowman, D. Koller, and L. F. Hodges. Travel in immersive virtual environments: An evaluation of viewpoint motion control techniques. In *Virtual Reality Annual International Symposium, 1997., IEEE 1997*, pages 45–52. IEEE, 1997.

[49] P. B. Pascoal, D. Mendes, D. Henriques, I. Trancoso, and A. Ferreira. Ls3d: Lego search combining speech and stereoscopic 3d. *International Journal of Creative Interfaces and Computer Graphics (IJCICG)*, 6(2):18–36, 2015.