Mid-Air Manipulation of 3D Models in (Semi-)Immersive Virtual Environments

Vasco Rodrigues
vasco.t.rodrigues@ist.utl.pt

Instituto Superior Técnico, Lisboa, Portugal

July 2015

Abstract

With the technological advances in recent years, there has been a growth of 3D objects manipulation techniques. The most notable examples are the mid-air interaction techniques, that allow users to manipulate objects in their surrounding space.

These techniques can differ in the freedom of manipulation they provide, the use of intrusive devices and the use of stereoscopy. They can also be used in different environments, such as immersive and semi-immersive virtual environments.

Recent studies concluded which techniques were the best in semi-immersive environments, regarding the time to complete a task. However these techniques can have a different performance when used in a different environment. This work had the objective of comparing techniques in both immersive/semi-imersive environments, to ultimately conclude which is the best regarding the time to complete a task, as well as the achieved precision.

With this work it was possible to observe that mid-air interactions in immersive systems are the most efficient and satisfying method for users. That is due to the direct manipulation of 6 degrees of freedom, which mimics the way with which we interact with physical objects.
Keywords: Manipulation, 3D Objects, Precision, Mid-Air

1. Introduction

In the recent years, the use of 3D objects has largely increased. They are used in several different areas like assembly lines, health software and architecture. With the surfacing of new technologies appears a sea of possibilities to interact with 3D objects, that differ from traditional approaches. One of these methods of interaction is the manipulation of 3D objects in mid-air, that shows a lot of potential and is worth of being explored. This could allow to achieve new solutions for object manipulation that are more natural to the users.

The currently used methods to interact with 3D objects make use of the mouse and the keyboard, preventing the users from exploring all of the objects' functionalities. These methods use the WIMP paradigm, (Windows, Icons, Menus, Pointers) while having a bi-dimensional visualization and preventing an interaction in a natural way.

The possibility of bringing the interactions with the 3D world to the space around the users, using 3D glasses and HMDs, allows to give an immersion feeling to the user. The 3D objects manipulation methods in the tridimensional space have been the subject of studies in recent years, with different approaches having been proposed. These differ in freedom of manipulation, the use of intrusive devices and also the use of stereoscopy. Some of these manipulation techniques were studied by Fonseca et al. [3] [8], with the purpose of finding which techniques were faster and more natural to users in a semi-imersive virtual environment.

With this work we wanted to study how different techniques work in immersive virtual environments and semi-imersive virtual environments, focusing on the time to complete a task and the achieved precision.

Basing our approach on the set of information found and the defined ideas for the study of non intrusive and natural object manipulation, we developed two systems to test two different manipulation techniques. Using a 3D television, a Kinect camera, a pair of 3D glasses, an HMD Oculus Rift and two Wiimote controllers with a Motion Plus adapter, the user can manipulate immersively and semi-immersively objects. The user interacts with the Wiimotes, viewing the scenario either through the TV and the 3D glasses or through the Oculus Rift. In these prototypes the user has two techniques to interact with the objects. The 6DOF

Hands technique allows a direct manipulation of the objects with each of the user's hands so he can move and rotate freely an object as in a real world interaction. There is also the possibility to scale objects using the free hand. The other technique is called Handle Bar and tries to simulate a barbecue spit. The user grabs a bar with both of his hands while the object stands between them. Then the user can manipulate the object by moving the bar like it would move on the real world.

To evaluate the implemented techniques we carried out user test sessions, where we compared both implemented techniques in two different environments, being one of them an immersive setup, and the other one being semi-immersive. The 6DOF Hands technique proved to provide better precision on the placement of the objects, as well as taking less time to do so. The immersive setup provided better visualization, making the tasks easier, thus being more pleasant to the users.

1.1. Contributions

During this work we created two applications, as a result of the need to comunicate between different computers. Each system was executed in different computers, as well as other devices. The application KinectDataReader sends the data capture by the Kinect camera to the system. MoveDataReader reads and sends to the system the buttons' statuses and the orientation of the Playstation Move controllers.

2. Related Work

Nowadays, the manipulation of 3D objects is used in several applications, being indispensable in several areas. We can easily find the adoption of 3D manipulations in many fields, such as design and assembly of components in engineering, videogames, the creation of architectural models, medical software, among others. With the great usefulness that the manipulation of tridimensional objects presents, it has been the subject o several studies and constant development, aided by the improvement and surfacing of new technologies, which turn the tasks easier to complete. Here we describe some of the works related with 3D object manipulation.

Regarding 3D object manipulation on multitouch surfaces, Cohé et al. [1] created tBox. It allows to manipulate objects with a 3D interaction, taking advantage of a direct and independent 9DOF control. The system allows very precise translations, using a cube whose edges contain the object to manipulate, as we can see in Figure 1. The object can be moved by dragging the edges and can be scaled by moving away or closer two parallel edges. To rotate the object, one can drag the cube's faces.

Focusing on gesture recognition, Wang et al. [13] created a system that tracks the hands and can recognize a set of gestures in 6DOF. The gestures are created as real world metaphors. It uses a snap system that allows the users to join objects with more precision. The implemented gestures allow to move or even throw an object, instead of dragging it. It is also possible to rotate the camera and the object. The great advantage of this system is that it captures the hand gestures without being intrusive, using two Kinect cameras above the interaction space (Figure 2). The background is removed, leaving just the hands, which are then compared to the estimated poses of the gestures on the database. However, it can only be used on a setting where the background can be removed, and with long sleeves.

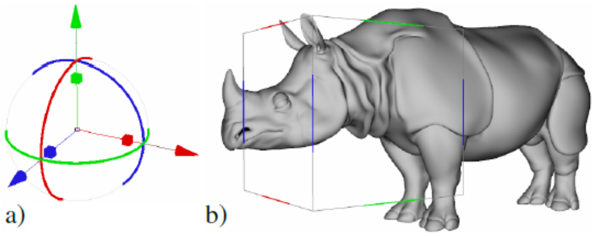

Figure 1: tBox (Cohé et al. [1]): a) Standard 3D transformation tool; b) tBox to control the object's 9 DOF.

Following the same line, Hilliges et al. [4] created Holodesk, which combines a transparent screen and a Kinect camera to create the illusion that the user is interacting directly with the virtual environment. The interaction space is located below the screen, giving the impression that the objects are next to the user's hands. The Kinect camera detects the hands and other physical objects' position, allowing to manipulate the virtual objects using them, as we can see in Figure 3.

De Araùjo et al. [2] developed Mockup Builder, a semi-immersive environment where it is possible to create and manipulate with gestures on and above a surface. It uses stereoscopy, adapting the user's perspective of the 3D environment according to his head position, captured with a Kinect camera. The hands' positions are detected with two Gametrack

Figure 2: System created by Wang et al. [13] recognizing a user's hands.

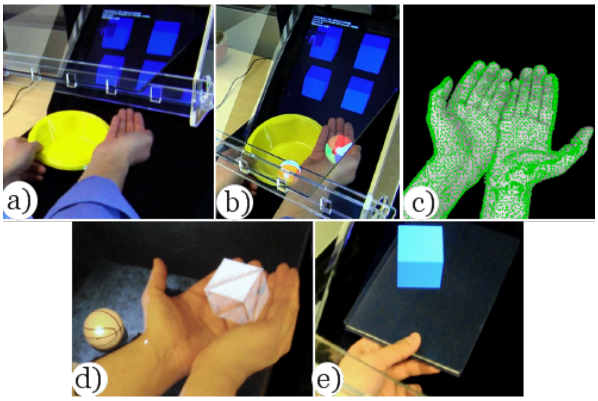

Figure 3: Hands and physical objects' interaction with virtual objects on Holodesk [4].

devices, allowing the user to manipulate objects with 7DOF. By combining the use of two hands with the flexibility of a continuous space, this system allows to change easily between 2D and 3D gestures. The user can choose the best type of manipulation according to his tasks.

Song et al. [11] proposed the barbecue spit metaphor to control virtual objects in space (Figure 4). They called it Handle Bar, and it tries to mimic a familiar situation to the users: the use of a barbecue spit, being held with both hands and maintaing an object on its center. This system allows to manipulate an object using 7DOF. The user holds an object putting each each hand next to the object, moving his hands to move it and rotating it by rotating his hands by the object's center. To scale the object, the user can move his hands away or closer to each other while holding the object. This solution offers a natural way to manipulate the objects while using a familiar situation to the users and being non intrusive.

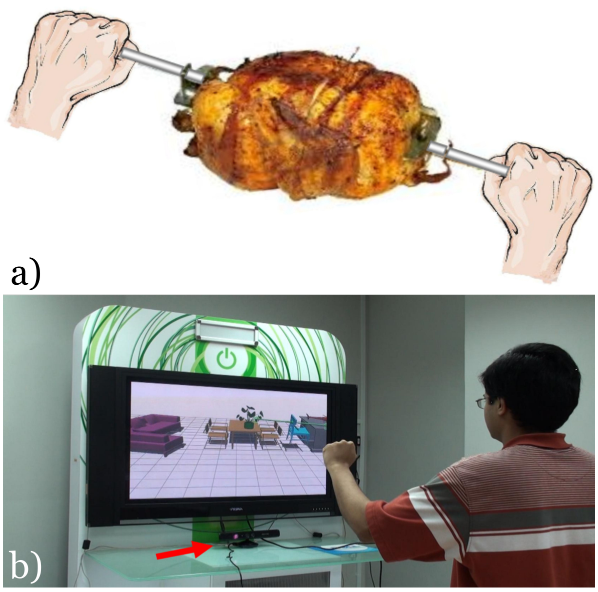

Figure 4: Barbecue spit metaphor used by Song et al. [11]: a) Metaphor example; b) User trying the metaphor on the system.

Still regarding 3D object manipulation, Fonseca et al. [3] [8] made a comparison between five different above and on the surface techniques (Figure 5). The system combines non intrusive motion tracking technology easily available with a multi-touch table. Using stereoscopy, four techniques are used above the surface, while the fifth uses touch gestures. Above the surface, the developed techniques were 6DOF Hand (simulates physical interactions and offers 3 DOF on translation and rotation, using the free hand to scale), 3DOF Hand (translation, rotation and scaling with both hands), Handle Bar (based on the work of Song et al. [11]) and Air TRS (grab with one hand to move, both to rotate and scale with the distance between both hands). On the surface, the TRS Touch + Widgets technique relies on touching the surface below the object and using widgets. The study concluded that direct interactions above the surface are efficient and satisfactory, while still having problems caused by the lack of precision and hand occlusion. 6DOF Hand was considered more natural to use for reproducing the direct interactions with physical objects, while Handle Bar was as fast as 6DOF Hand.

3. Research Study

We wanted the user to be able to make air interactions with the system we wanted to develop. For that it was necessary that the users gestures were recognized, by detecting his hands position and orientation. With this in mind, we started researching and experimenting, trying to find the best way of obtaining the needed data.

3.1. Motion Capture

The first problem was how to create a setup that was able to capture the users gestures and position.

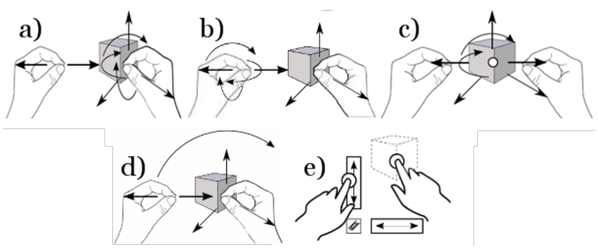

Figure 5: Techniques on the study by Fonseca et al. [3] [8]: a) 6DOF Hand; b) 3DOF Hand; c) Handle Bar; d) Air TRS and e) TRS Touch + Widgets.

This would allow to adapt his perspective of the virtual world. Different options were studied and tested, allowing us to find the most adequate.

The first approach was to try to recreate the system used by Mendes et al. [8], as it was an accessible system that allowed the techniques that would be implemented to function correctly. The system uses the Nimble VR software, by Wang and Twigg [12], and a depth camera above the users hands. It allows to capture several points on the users hands and detecting gestures. However, due to many updates since its use by Mendes et al. [8], it can no longer recognize the hands position correctly.

The next hypothesis was the use of the Leap Motion system, by Holz [5]. It allows to capture the hands movements with great precision using inexpensive hardware. However, its interaction area is limited to a 61cm radius from the device, which is smaller than the area needed for our system. Also, it can't follow the hands' rotation when they are closed.

Another studied option was the use of the Optitrack system, by NaturalPoint. Using 10 cameras to capture the markers position, these could be mapped virtually. It was a reliable solution, however it defeated our initial purpose of having a system accessible to the common user, as each camera as cost of €500.

With this problem in mind, we opted to use the Microsoft Kinect 2 camera to perform our manipulation techniques. This solution can detect the users body positions reliably, offering a wide interaction area. However, despite its reliability in capturing positions, we came to the conclusion that the hand orientation detection fails.

3.2. Detecting Hands' Orientation

With the surfacing of this problem, we needed to research other technologies that could be used with the Kinect camera and still capture the movements and orientation of the hands. Several solutions were tested.

The first solution was to try and explore the potentialities of the Wiimote Plus controllers. These possess a gyroscope, allowing to obtain the hands orientation while interacting with the objects using the buttons. After acquiring the controllers, several libraries such as Wiiuse [7], WiimoteLib [9] and GlovePie [6] were explored, trying to obtain data from the controllers. However none of them is compatible with the Plus version of the Wiimote, returning blank or wrong data.

It was also explored the idea of using a smartphone that had a gyroscope, straped to each hand. The user could touch the touchscreen of the phone when he wanted to interact with the objects or just make the gestures to manipulate them. However, the models that have gyroscope are too expensive (and two would be needed) and impractical, due to their larger size.

Following the controllers with gyroscope idea, we tried to use two PlayStation Move controllers. Using the work of Perl et al. [10] as a basis, we were able to integrate them on the Project. This solution allowed to get the interaction with the objects, using its buttons, and their rotation. Their silicone sphere also allowed to provide visual feedback to the users. In spite of the success in integrating the controllers, this solution wasnt the ideal. The orientation wasnt always reliable, with the gyroscopes signaling rotations when the controllers were idle.

The last solution was to use Wiimote controllers and attach a Motion Plus adapter. This provides the controllers with the same functionalities as the Wiimote Plus controller. Contrary to the Wiimote Plus, the use of the adapter was able to work with the GlovePie library. The gyroscopes data was received correctly, allowing to recreated the correct controller movements virtually.

4. Implementation

The objective of this work was to study the behaviour of mid-air manipulation techniques in imersive and semi-imersive virtual environments. We start by presenting the implemented system used to test the interaction techniques. We also present the implemented manipulation techniques that allow a simple direct interaction, as well as solutions to control the camera on the 3D environment.

4.1. Hardware Configuration

The developed solution uses a Microsoft Kinect 2 camera to track the movement of the user's hands and head. In order to interact with the 3D objects, the user has two Wiimote controllers with a Motion Plus adapter. This wireless device, usually used to play games, has an adapter that allows to obtain the user's hands' orientation. The buttons of the controller allow the user to use the manipulation techniques to interact with the objects.

Another device used with our solution is the HMD Oculus Rift 2, that allows an immersive visualization of the environment. With the help of the Kinect camera, it's possible to detect if the user has changed his position and/or his orientation, updating his perspective and position in the virtual world.

As for the stereoscopic visualization, a combination of a Samsung 3D television and a pair of Samsung SSG-5100GB 3D glasses is used. This setup allows a semi-immersive visualization above the television screen. It's possible to view the used devices in Figure 6.

4.2. Architecture

With this work we wanted to create an environment that would allow to find more natural tridi-

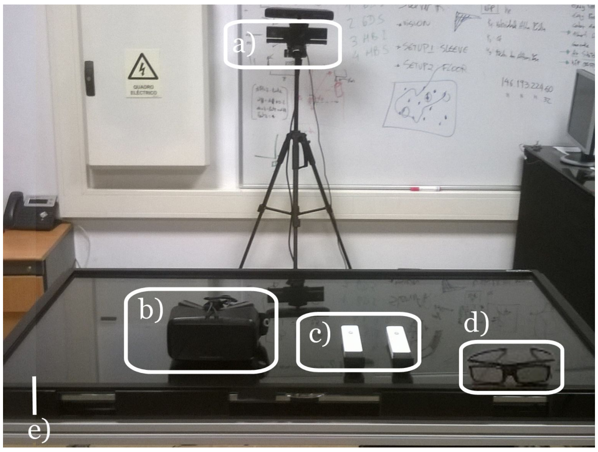

Figure 6: Used devices with the system: a) Kinect 2; b) Oculus Rift 2; c) Wiimote controllers; d) Samsung SSG-5100GB 3D glasses and e) 3D television.

mensional manipulations. With this goal in mind, we developed the architecture represented in Figure 7. The system consists in five different modules, with two of them being the core of the architecture.

4.2.1 Gestures Module

The Gestures Module captures the user's movements using a Kinect 2 camera. This data is sent to the module through UDP, which then maps it to the virtual world, recreating the movements virtually in real time. The data contains the user's hands' position.

4.2.2 Interaction Module

The Interaction Module is a pillar of the system. This module allows to use the implemented manipulation techniques. The data received from the Kinect and the Wiimote controllers from the Gestures Module include the buttons' statuses as well as their orientation, given by the gyroscopes in the Motion Plus adapters, and position. This data allows to recreate the manipulation techniques, grabbing the objects to move, rotate and scale them.

4.2.3 Scene Module

The Scene Module is responsible for dealing with the logic that allows to represent any element in the 3D space, as well as its movements. It allows to show the result of moving an object, rotate it and scale it. It also allows to move through a scene with ease and show different initial states for each. It communicates directly with the Render Module, mapping the visualization in Oculus Rift and the stereoscopic surface.

4.2.4 Viewpoint Module and Render Module

The Viewpoint Module uses the Kinect data about the user's head position to calculate its position on the virtual world. These calculations are sent to the Render Module, which can adapt what is seen by the user on the virtual scene.

In this manner, the Render Module shows the user the virtual environment where he interacts, using Unity 3D to create the virtual scenarios.

4.3. Interaction Techniques

In this work we implemented two interaction techniques which allow the user to manipulare 3D virtual objects, using stereoscopy or virtual reality. The interaction is made in the space in front of the user. Based on the works of Wang et al. [13], Song et al. [11] and Fonseca et al. [3] [8], the chosen techniques were 6DOF Hands and Handle Bar. These techniques allow a familiar manipulation interaction, having better performance than other studied techniques.

These techniques allow the user to interact in 7 DOF: 3 DOF to translate, 3 DOF to rotate and 1 DOF to scale. In each technique the user can choose which hand to use as dominant to manipulate the object. Since that for each technique the user needs both hands to manipulate in 7 DOF, only one object can be manipulated at a time. We present these techniques with more detail on the next sections.

4.3.1 6DOF Hands

In 6DOF Hands, the user's hands are tracked in real time, allowing to simulate the way one holds, moves and rotates an object with one hand in the physical world. As one can see by the name, one hand allows 6 DOF, three to move and three to rotate. This technique is based mainly on the work of Wang et al. [13].

The user touches with his Wiimote controller the object he wants to interact with, holding it with the B button. Then, the object is moved by following the movements of the hand that holds it. The rotation is detected by the controller, as the user rotates his wrist.

To scale the object, the user uses his other controller. The hand does not need to be touching the object. The scaling is indirect, using the distance between the hands and mapping hands being closer or further away with reducing or enlarging the object's size. We can see an interaction using 6DOF Hands in Figure 8.

4.3.2 Handle Bar

Based on the work of Song et al. [11], we adopted the Handle Bar metaphor. This simulates the use of

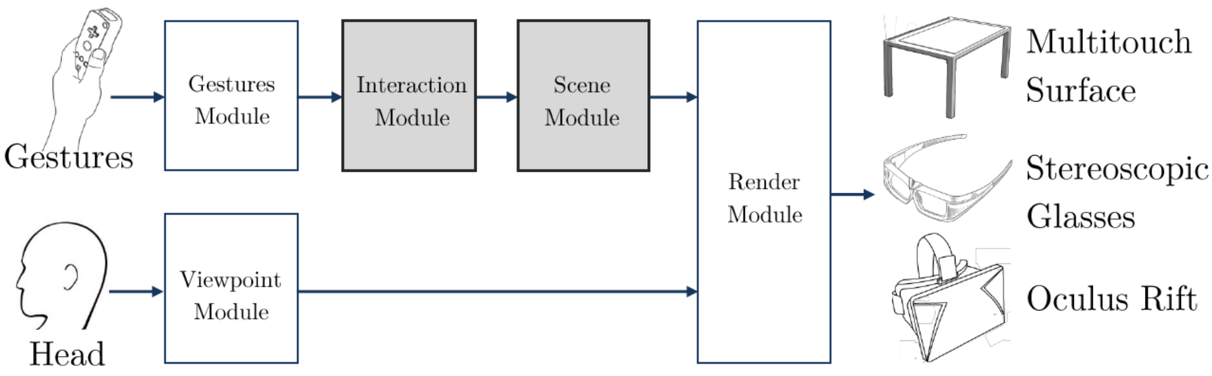

Figure 7: System's architecture.

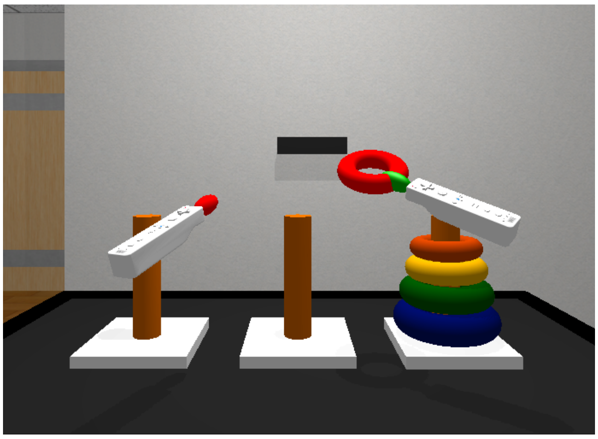

Figure 8: Manipulation using 6DOF Hands.

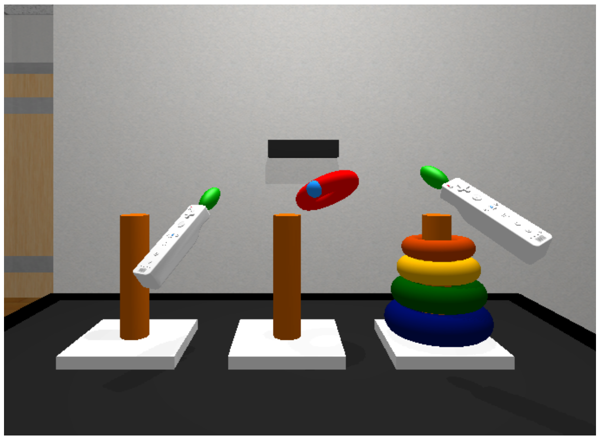

Figure 9: Manipulation using Handle Bar.

a barbecue spit to grill chicken. The user holds the spit with both hands, whereas the chicken is placed at its midpoint.

The user can grab and manipulate objects using the midpoint of his hands. This point is virtually represented as a blue dot. To hold the object, the user puts this dot on it, grabbing with both hands. With the object being held like a bar is passing through it, the user can manipulate it. He can move and rotate the object, which will mimic the the bar movements. To scale the object, the distance between both hands is used. We can see an interaction using Handle Bar in Figure 9.

5. Results

In order to compare the implemented techniques, it was necessary to conduct a set of user tests. The techniques were evaluated in regards to ease of usage, precision and time. Two systems were developed to allow the comparison immersively and semi-immersively. Each user performed three placement tasks with each technique in each system.

5.1. Systems

Two different systems were created to test the implemented techniques. One represents an immersive virtual environment (IVE), and the other a semi-immersive virtual environment (SIVE). Each system has a collisions mechanism that detects when a controller is touching a graspable object. The user must touch the objects with an ellipsoid placed in front of the controllers in order to grab them. When the ellipsoid is touching a graspable object, a green wireframe appears surrounding the object.

None of the techniques has physics, to allow a completely free manipulation, which also helps in measuring the precision of the tasks. After each task, the relevant data is automatically saved to a file, allowing its review later.

5.1.1 Immersive Virtual Environment - IVE

The IVE system is a virtual recreation of the room where the tests take place. The goal is to test manipulations with 3D objects, giving the illusion that the user is inside the virtual world.

The user is able to look in each direction, with the help of the orientation given by the Oculus Rift HMD. He can also move on the scene, as the Kinect detects his head's position and applies the needed transformations. This gives freedom to his movements. The user's arms are also recreated using the position data from the Kinect, in an attempt to give a bigger sense of presence in the virtual world. In Figure 10 we can see virtual environment of the IVE system.

5.1.2 Semi-Immersive Virtual Environment - SIVE

The environment in the SIVE system consists of a plane that overlaps the 3D television's surface. This allows to use stereoscopy, giving the illusion to the user that the objects are above the surface.

By seeing the objects above the surface, the user can interact with them using the Wiimote controllers. The user's perspective of the scene is given by his head's position with the help of the Kinect camera. In Figure 11 we can see virtual environment of the SIVE system.

5.2. Methodology

The tests were structured accordingly to the following scheme:

- **Introduction**

  When starting a new user test, each system was presented, as well as its capacities. It was also explained how to intereact with each prototype and the technological limitations present in monitoring the user's hands.

- **Training Session**

  It was shown to the user how to move, rotate and scale objects with each technique in each system. We allowed the user to understand how to interact and train his interactions.

- **Tasks Execution**

  We randomly determined which system and technique was the first to be used by the user, to guarantee that the order of usage did not influence the results. It was asked to each user to execute a set of three equal tasks with both techniques in each system. The order of the tasks was predetermined, with an increasing difficulty.

- **Final Questionnaire**

  After the test sessions, the participants filled in a questionnaire to evaluate different aspects on each technique and system, such as ease of usage, preference and fun factor.

- **Profile Questionnaire**

  Finally, the users filled in a questionnaire with the objective of getting to know their profile: age and experience towards stereoscopy and other devices.

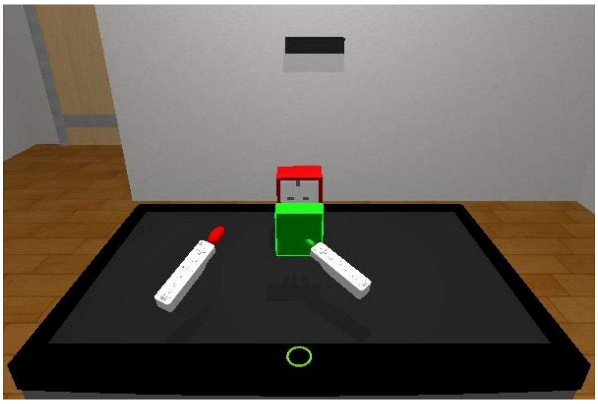

Figure 10: IVE system.

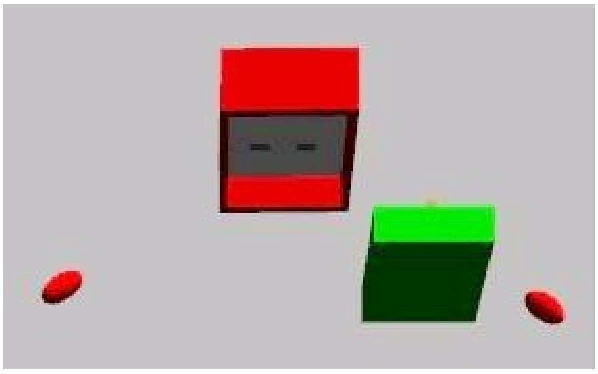

Figure 11: SIVE System.

5.3. Tasks

Knowing that solids are present in one's life early, it was decided that the tasks would follow the idea of fitting objects in holes with its shape. The tasks had incremental difficulty, starting with a simple task and ending with one that needs a better control of the technique to be concluded. As a result, there was no gravity nor physical collision between objects. Also, the user can only manipulate the object that will be fit in the hole. The tasks were the same in each technique-system pair.

The users had a training session for each technique-system. In the scenario the user had a replica of the Tower of Hanoi game, as we can see in Figure 12. With this, the user could manipulate the rings freely and learn how to interact in each technique and system.

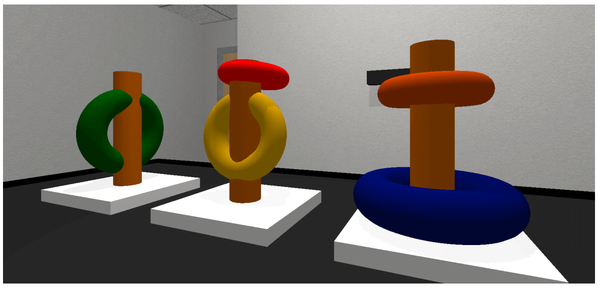

Figure 12: Task performed during the training session.

During the tests, the tasks consisted in fitting the green plug in the red socket, with the highest possible precision. Only the green plug can be manipulated. In each system the tasks were executed by incremental difficulty. On the first task, the user moved the plug, on the second task the user needed to move and rotate it, and on the third the user needed to move, rotate and scale the plug to achieve the final position. In Figure 13 we can see the initial state of the tasks.

5.4. Apparatus and Participants

The test sessions were performed on a closed environment, without the possibility of external influences. It was used a Kinect 2 camera, two Wiimote controllers with Motion Plus adapters, 3D Samsung glasses and the HMD Oculus Rift. Both computers used by each system had a Windows 7 operating system, with an Intel Core i7 - 3770K 3.5 GHz CPU and 16 GB RAM.

The evaluation had 20 users, one female and nineteen males, with ages between 19 and 35. Only four users hadn't tried stereoscopy, and eight never had experienced virtual reality. Nineteen of the users had already played with a gaming console, with ten having used the Kinect, fourteen the Wiimote and eight the Playstation Move. On Figure 14 we can see users interacting with each system.

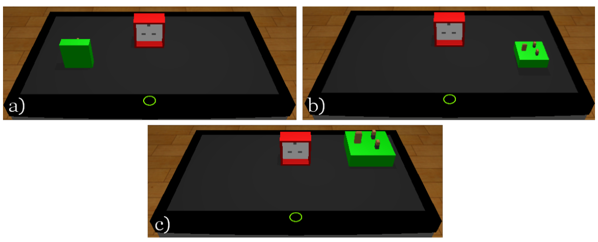

Figure 13: Initial state of the tasks: a) Task 1; b) Task 2; c) Task 3.

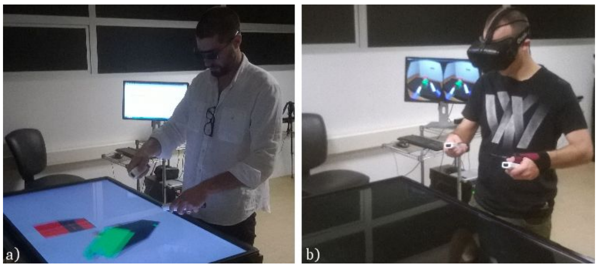

Figure 14: Users testing the systems: a) SIVE and b) IVE.

5.5. Results and Discussion

This evaluation allowed us to understand the technique that felt most natural to the users. We analysed our measures using the Wilcoxon Signed-Rank Test, comparing precision and time spent by each user to conclude a task.

5.5.1 Task 1

The difference from the ideal angle of the plug was smaller in 6DOF Hands - IVE than on Handle Bar - IVE ($Z = -2.725$, $p = 0.006$), as it was easier to rotate an object in 6DOF Hands. Handle Bar has an axis which can't rotate. The time to complete the task was smaller on 6DOF Hands - IVE than on 6DOF Hands - SIVE ($Z = -2.763$, $p = 0.006$). With the ability to see many different angles using the Oculus Rift, it's faster to place an object.

5.5.2 Task 2

The distance from the ideal position of the plug was less with 6DOF Hands - IVE than with Handle Bar - IVE ($Z = -2.576$, $p = 0.01$). That might be caused by the need of rotating the plug, and the lack of a rotation in Handle Bar might tire and bore the users, making them placing the plug with less accuracy. The difference from the ideal angle of the plug was smaller in 6DOF Hands - IVE than in Handle Bar - IVE ($Z = -2.978$, $p = 0.003$), being also smaller in 6DOF Hands - SIVE than in Handle Bar - SIVE ($Z = -2.203$, $p = 0.028$). The task was completed faster with 6DOF Hands - IVE than with Handle Bar - IVE ($Z = -2.800$, $p = 0.005$) and 6DOF Hands - SIVE than with Handle Bar - SIVE ($Z = -3.360$, $p = 0.001$).

5.5.3 Task 3

The distance from the ideal angle of the plug was less with 6DOF Hands - SIVE than with Handle Bar - SIVE ($Z = -2.016$, $p = 0.044$). The rotation difficulty with Hanfle Bar is even more noticed in the semi-immersive system, as it's more difficult to

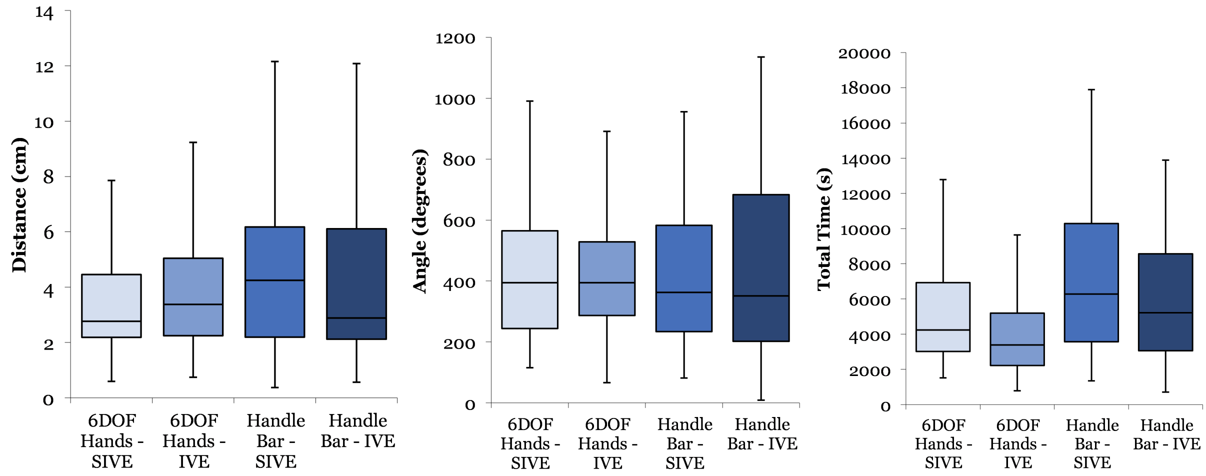

Figure 15: Measures obtained in the tasks.

the user to see the objects from every angle, as well as understand their true position. Being this the third task of the users with each technique-system pair, the users were able to fit the plug with ease in the same position.

5.5.4 Qualitative Analysis

On the final questionnaire, the users were asked to classify their experience with each technique and system. The users thought that learning how to interact with 6DOF Hands - IVE was easier than with Handle Bar - IVE, 6DOF - SIVE and Handle Bar - SIVE, globally ($Z = -3.220$, $p = 0.001$ and $Z = -3.066$, $p = 0.002$ and $Z = -3.440$, $p = 0.001$) and in the rotation ($Z = -3.779$, $p < 0.001$ and $Z = -3.542$, $p < 0.001$ and $Z = -3.495$, $p < 0.001$). Also in the rotation, 6DOF Hands - SIVE was easier than Handle Bar - IVE ($Z = -2.097$, $p = 0.036$). To translate, 6DOF Hands - IVE was easier than 6DOF Hands - SIVE and Handle Bar - SIVE ($Z = -3.116$, $p = 0.002$ and $Z = -2.799$, $p = 0.005$). Scaling was easier with 6DOF - IVE and Handle Bar - IVE than with Handle Bar - SIVE ($Z = -2.435$, $p = 0.015$ and $Z = -2.516$, $p = 0.012$). The preference of the techniques on the immersive system, specially 6DOF Hands, is due to its simplicity and ease of observation with Oculus Rift, as well as a bigger precision. On the 3D television's surface the objects could disappear from the screen with the side movements of the user, as the system tried to adapt the user's perspective of the objects according to his head's position.

The users preferred the immersive techniques, the favorite being 6DOF Hands - IVE, against Handle Bar - IVE, 6DOF - SIVE and Handle Bar - SIVE ($Z = -3.036$, $p = 0.002$ and $Z = -3.294$, $p = 0.001$ and $Z = -3.587$, $p < 0.001$). They also preferred Handle Bar - IVE against Handle Bar - SIVE ($Z = -2.353$, $p = 0.019$). The fun factor got similar results, with the users having more fun in 6DOF Hands - IVE, than in Handle Bar - IVE, 6DOF - SIVE and Handle Bar - SIVE ($Z = -3.082$, $p = 0.002$ and $Z = -3.800$, $p < 0.001$ and $Z = -3.970$, $p < 0.001$), and in Handle Bar - IVE than in Handle Bar - SIVE ($Z = -3.080$, $p = 0.002$). These preferences are in the same line as the interaction preferences.

Finally, comparing the systems, immersive and semi-immersive, the users felt more present in the immersive system, with that feeling being smaller in the semi-immersive system ($Z = -3.775$, $p < 0.001$). It is possible to observe each object from every different angle immersively, whereas semi-immersively the television's edges limited the visualization. To most of the users (seventeen), being able to see their arms on the immersive system helped them feeling more present on the virtual environment.

6. Conclusions

The constant technological advances have allowed the creation of more and more scenarios and environments that offer unique interaction capabilities. In recent years the use of 3D objects has grown. These are used in many different fields, such as assembly lines, health and architecture software, among others. The most recent manipulation techniques allow the interaction with 3D objects in mid-air, providing a more natural and free manipulation.

In this work we tried to identify natural interaction techniques, which had to use gestures familiar to the users. After a study about existing techniques and works about manipulation on the tridimensional space, two techniques were chosen. The chosen techniques were 6DOF Hands and Handle Bar. These were implemented in two different environments, one immersive (IVE) and the other semi-immersive (SIVE). In this work we compared the behaviour of these techniques on both environments to find out which one is the best in each environment, regarding time to complete the task and the achieved precision.

To evaluate each technique in each environment, we carried out user test sessions. The users were asked to test each technique-system pair by performing three tasks of incremental difficulty. In each task the user had to fit a plug in a socket. The 6DOF Hands technique had higher precision and smaller task completion time, specially when rotations are needed. Handle Bar had a major constraint by not allowing to rotate in one axis. With the surpass of this problem, it may be possible to get better results. The completion time was also smaller in the IVE system. That can be explained by the fact that it's easier to observe objects from different angles and interact with them immersively. However the results of both techniques can still be improved, due to the lack of millimetric precision, essential in different fields.

The users felt it was easier to interact (move, rotate, scale) using the 6DOF Hands technique on the IVE system. This was due to its simplicity and ease of observation, existing also a higher level of precision, as mentioned before. This was also the users' preferred technique, being also regarded as the most fun to use. The techniques on the IVE system were considered easier to use, since on the SIVE system it was more difficult to view the objects.

Still, we believe it is possible to improve and complete this work. The implemented techniques for 3D object manipulation can be improved and there's also room to study and develop new interaction techniques, that can achieve greater precision than the tested ones, trying however to maintain the same level of familiarity to the users. It would also be relevant to try to combine new technologies that could allow to obtain the user's hands' orientation, without resorting to holding objects on their hands. With the gathered knowledge and conclusions obtained in this work, we created a foundation to support future works on this topic.

References

[1] Aurélie Cohé, Fabrice Dècle, and Martin Hachet. tbox: A 3d transformation widget designed for touch-screens. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2011.

[2] Bruno R. De Araùjo, Géry Casiez, and Joaquim A. Jorge. Mockup builder: Direct 3d modeling on and above the surface in a continuous interaction space. In *Proceedings of Graphics Interface 2012*, 2012.

[3] Fernando Fonseca, Daniel Mendes, Bruno Araùjo, Alfredo Ferreira, and Joaquim Jorge. 3d mid-air manipulation techniques above stereoscopic tabletops. ISIS3D: Interactive Surfaces for Interaction with Stereoscopic 3D - ITS 2013, 2013.

[4] Otmar Hilliges, David Kim, Shahram Izadi, Malte Weiss, and Andrew Wilson. Holodesk: Direct 3d interactions with a situated see-through display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2012.

[5] David Holz. Leap motion. `developer.leapmotion.com`, Outubro 2012. Acedido em: 10-2014.

[6] Carl Kenner. Glovepie 0.45. `glovepie.org`, Dezembro 2011. Acedido em: 13-04-2015.

[7] Michael Laforest. Wiiuse. `sourceforge.net/projects/wiiuse`, Abril 2013. Acedido em: 03-2014.

[8] Daniel Mendes, Fernando Fonseca, Bruno Araùjo, Alfredo Ferreira, and Joaquim Jorge. Mid-air interactions above stereoscopic interactive tables. 2014.

[9] Brian Peek. Wiiuse. `brianpeek.com/page/wiimotelib`, Janeiro 2009. Acedido em: 03-2014.

[10] Thomas Perl, Benjamin Venditti, and Hannes Kaufmann. Ps move api: A cross-platform 6dof tracking framework. In *Proceedings of the Workshop on Off-The-Shelf Virtual Reality.*

[11] Peng Song, Wooi Boon Goh, William Hutama, Chi-Wing Fu, and Xiaopei 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*, 2012.

[12] Robert Wang and Chris Twigg. Nimble vr sdk software. `nimblevr.com/download.html`, Julho 2012. Acedido em: 10-2014.

[13] Robert Wang, Sylvain Paris, and Jovan Popović. 6d hands: Markerless hand-tracking for computer aided design. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, 2011.