The Blind Men and the Elephant - Enhancing Virtual Reality Interactions

First Author
Taraburcă Radu
whx862

Abstract
In this paper, we are exploring the ways to enhance the interactions with the virtual reality environment. We investigated the possibility of improving the experience by interacting with a manufactured object. We determined what type of manufactured object can give a more authentic feeling and which is also efficient from the point of view of occupied space. In this work, we proposed 2 types of manufactured objects. An object that can change its shape and an object that is an amalgam of common geometric shapes that we can find in everyday objects. Besides that, we created a 3D environment in VR where we can interact with these objects. To test these objects and see how viable they are, we prepared a series of experiments. Based on the results obtained, we want to help the development of interaction systems with the virtual environment that improves the feeling of immersion but also does not take up a lot of space.

Author Keywords
Virtual Reality (VR), Tangible Objects, 3D Printing, Robotic Arm,Grasp Identification

CCS Concepts
•Human-centered computing $\rightarrow$ **Human computer interaction (HCI);**

CHI'20,, April 25–30, 2020, Honolulu, HI, USA
ACM 978-1-4503-6819-3/20/04.
https://doi.org/10.1145/3334480.XXXXXXX

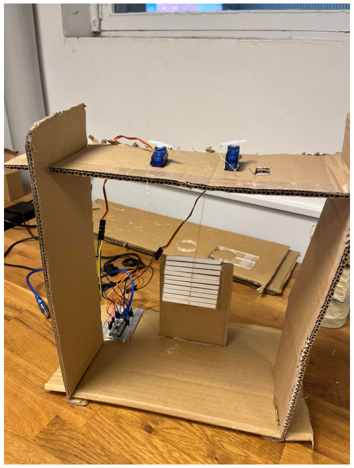

Figure 1: Shapeshifting object last iteration.

Figure 2: Commun feature object coated.

Introduction

Virtual reality (VR) technology offers immersive experiences, but the lack of physical touch can limit the realism and engagement of users. To address this limitation, incorporating tangible objects into VR environments has shown the potential in enhancing the overall experience. However, creating and managing multiple 3D-printed objects for each virtual environment can be time-consuming and requires significant storage space. This project aims to improve the efficiency and practicality of incorporating tangible objects into VR experiences.

The problem addressed in this project is the development of an object that can improve the experience of immersion but which is also easy to manufacture and store. Therefore, we proposed 2 prototypes, an amalgam of fused geometric shapes, which is a simple object to manufacture but which is not very efficient from the point of view of space, and an object that can change its shape, which is more complicated to manufacture, but which is simpler to store. Besides that, we designed a system that is able to detect which object would be the most likely to be grabbed. We used several artificial intelligence models to achive it. This approach can be improved by adding grasp type detection using GrabNET[5]. We were inspired by other studies such as "Snake Charmer: Physically Enabling Virtual Objects"[1] and "Annexing Reality: Enabling Opportunistic Use of Everyday Objects as Tangible Proxies in Augmented Reality"[3].

Contributions

At the moment, there are few methods and objects that can increase the realism in the VR environment. So that our work contributes to the presentation of methods of manufacturing objects that can solve this problem. Besides that, we also contribute with a method to detect the objects that are most likely to be grabbed in the vr environment. And besides that, we contribute with information related to directions that we have not explored and with suggestions related to these future explorations that can contribute to future studies about interaction systems in the VR environment.

Related Work

Virtual Reality (VR) is a relatively new and rapidly evolving evolving field that aims to create immersive digital environments that simulate real-world experiences. As VR technology advances,and it achieving higher levels of realism, with scenes and object close to the clarity of the physical world, the lack of a tactile stimulus, to increase its level of immersion, is felt. Thus, the development of an object with which we can interact has become important. During the research and development process, a series of studies attract our attention.

Haptics & Physical Proxies

The paper "Snake Charmer: Physically Enabling Virtual Objects"[1] had a major impact on our idea, this being a starting point in our research. This research explores the possibilities of physically enabling virtual objects and it focus on simulating multiple sensations in an immersive environment using a robotic arm. The robotic arm can pick up and position physical objects and textures that serve as a proxy for the virtual content. In our study, the aspects we relied on from this work were those related to the robotic arm as a means of connection between the virtual and the physical world. Since in this work there was a greater focus on the texture of the objects, than on their shape, we decided to explore the latter.

Another work that influenced the way we thought and manufactured the proxy objects, was the work "Annexing Re-

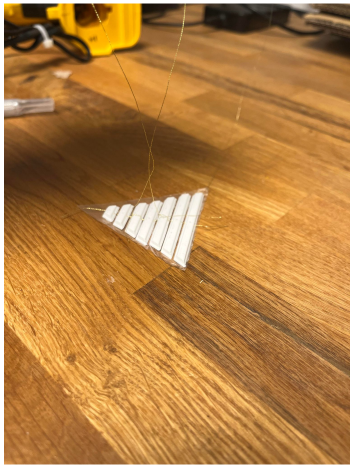

Figure 3: Shapeshifting object with strings.

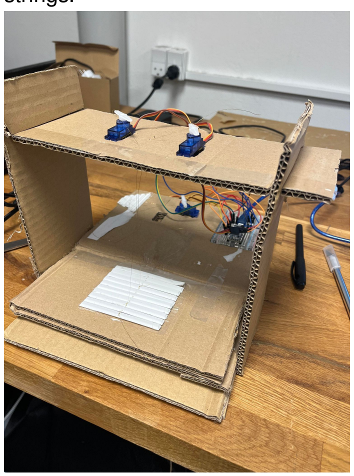

Figure 4: Shape shifting object iteration 2.

ality: Enabling Opportunistic Use of Everyday Objects as Tangible Proxies in Augmented Reality"[3]. This paper explores the implementation of a system that opportunistically annexes physical objects, from the user's current environment, besides that, it also talks about the design, its potential applications and benefits. This work influenced ours in connection with the use of shapes that resemble each other. We relied on this idea of similarity between primary geometric shapes when we designed one of our proxies.

Grasping and grasping in VR

When we talk about the interaction with an object, whether in the virtual or physical environment, we cannot help but talk about grasping the object. Grasping is a complex human mechanism to interact with an object. Depending on how we want to use a certain object, we can have a different grip, but not only that, this grip can also differ from one person to another [2]. In the virtual environment, grabbing objects is one of the main ways to interact with them. Considering that there are so many different types of grasping objects and because there are several types of interacting with objects in VR, we decided to focus on the pincer grasp, which is both a technique for grasping an object and a type of interaction in the vr environment with objects.

Our approach

The development process began by identifying and dividing the most important steps into stages. So this whole process was divided into several categories:

- Prototype design
- Development of prototypes
- Development of the 3d environment
- Development of a system for detecting the object to be grabbed
- Combining all these stages in a demo

Design process

To create a good design, we started by studying several works[[1], [3]] and what approaches their authors had. After that we did a brainstorming session and decided which prototypes we could have. We decided to focus on two prototypes. We chose to make two because one is more difficult to manufacture, and in case we fail to manufacture it we still have a functional solution.

Design and development of prototypes

We decided to design 2 types of objects (Fig 1 and Fig 2). An object that contains several common geometric shapes(Fig 2) that have been merged into a single voluminous object and a prototype of an object that can change its shape. The first object is direct, it is not the most efficient from the point of view of space, and it clearly cannot replicate specific forms that are not found in its composition, but in most cases it can be enough. When we designed this object, we had in mind the idea from the paper "Annexing Reality: Enabling Opportunistic Use of Everyday Objects as Tangible Proxies in Augmented Reality"[3], and how the chosen proxy objects are primary geometric figures. Thus, adapting the idea of primary proxy objects, we made the decision to manufacture a single proxy object with these properties, in order to reduce the space required to store them. When we say common objects, we are referring to the primary geometric figures, because these shapes can be found to some extent in any common object. Initially, this object was supposed to be 3D printed, but for reasons of efficiency of development, we decided to make it from cardboard by laser cutting several layers of each shape. I glued

Figure 5: Commun feature object.

Figure 6: Shape shifting object prototypes.

these pieces together one after the other and covered the final object with a layer of paper tape. We decided to cover it so that the user does not focus on the texture but only on the shape.

The object that changes its shape is based on the idea of caterpillars(Fig 1). By caterpillars, we mean the tracks of a tank. Just as the tracks from the tank are molded on any surface, this object is also able to change its shape. This object was originally thought to be 3D printed in its entirety, but later I made some changes to the initial idea. It has some boards that imitate teeth, and through which there are 2 holes. One of the holes crosses the entire width of the plate, and the other hole is perpendicular to the smaller surface of the plate. These boards are pierced by a wire(Fig 4) (F1) that passes through the first described hole. And the second hole is tied another wire (F2). The F1 thread has the role of main support and to produce the necessary tension to join all the face plates, while the F2 threads have the role of more precisely manipulating the desired shape or creating bumps. These plates are glued to a piece of plastic bag, which acts as a membrane. We made the decision to use a plastic bag instead of printing a back on which to attach the plates, because it is much more pliable than hard plastic, has less resistance when a plastic deformation force is applied and will not suffers a visible deformation after being bent. As a better alternative to the plastic membrane, we can use the technique described in paper "Stretching the Bounds of 3D Printing with Embedded Textiles"[4], where 3D objects are printed directly on fabrics. Thus, this approach reduces the manufacturing time of the object that changes its shape because it eliminates the step of manually gluing each plate on the membrane and reduces human error in relation to the distance between the plates. But besides that, the textile material is more durable than the plastic membrane and is more resistant to wear.

The wires are connected to some motors that will be set in motion by an ItsyBitsy(Fig 10). The ItsyBitsy will receive a notification with the object to be grabbed and will manipulate the threads in such a way as to change the shape of the object. These changes are made programmatically, according to some predefined schemes of pulling the corresponding wires.

Initially the object was thought to be made as an icosahedron where all the mechanical parts were inside, but later we changed the design to reduce the manufacturing complexity. We chose the icochasedeon because it was a geometric figure similar to a sphere and we wanted to go with the idea of a blind man changing his shape. This idea involved having the 20 faces (Fig 6) in the shape of a triangle in which we had printed these jagged plates. In the second iteration of the design(Fig 4), we decided to make the object as a straight plane. Thus we can upscale the object as much as needed. The object was then incorporated into a structure to be manipulated. It was fixed at one end and placed on a flat surface. The structure has a frame that comes on top of the object. On this frame, we fixed the servomotors for the fine manipulation of the plates through the F2 wires and on the flat surface where the object was fixed is the servomotor responsible for the manipulation of the main wire F1. In the third iteration of the object, we decided to change the design of the structure in which the object was embedded, to make the interaction easier. The new manipulation is done in the air, the object keeping its semi-fixed state. The new structure has the same frame but much higher and in the area where the object was fixed to the plane, we added a perpendicular plane, to which I fixed the object. So the user can interact with the object much more freely and the object can change its shape more easily.

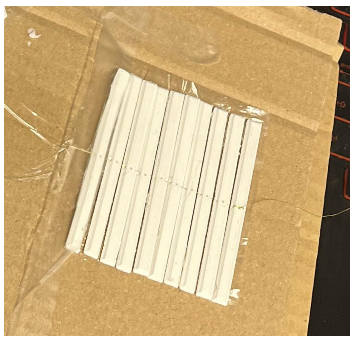

Figure 7: Shape shifting object prototypes last iteration.

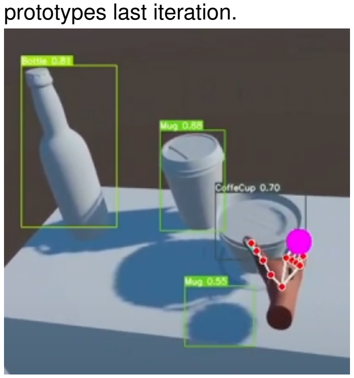

Figure 8: Object and hand detection in VR.

The virtual environment

The virtual reality environment(Fig 9) was created in Unity. To achieve this, and for a fast and efficient development, we used the XR Interaction Toolkit plug-in. It provided a wide range of objects and scenes. To create the virtual environment, we took a prefabricated scene and updated it. In addition to the objects provided by the plug-in, we added common objects such as a can, a bottle, a coffee cup, and a glass. In total, there are 4 platforms on which these objects can be found. On one pedestal we find geometric figures, on the second we find candles of different sizes and a lighter, on the third we find a hammer and a sprinkler and on the last, we find the objects listed previously. These pedestals are placed at a distance from each other in order to leave enough processing time to detect the object that will most likely be grabbed. For some of the objects, it was necessary to add a new anchor to make their grasping more natural.

Object detection

The object closest to the hand will most likely be grabbed. To detect which is the closest object(Fig 8) to the hand, we used some artificial intelligence models. For object detection, we used a custom version of the YOLO V7 model, and for hands tracking, we used the TensorFlow hand detection model. To predict which is the most likely object to be grabbed, we calculate the distance between the coordinates of the hand and the coordinates of the objects seen. And then we make a comparison between the resulting distances to determine which is the closest object. To train the YOLOV7 custom model, we captured images of each object in the vr environment and created a data set of approximately 5000 images. To collect the 5000 images, we created a python program that captures 20 frames per second and we used the 3D viewer application from Windows to view the object. With the help of the program made in python, we selected an area on the screen where to capture the object, and in 3D view we moved the object in 360 degrees on all 3 coordinate axes. In addition to this, we did manual labeling on each image to have the exact area where the desired object or objects are located. We did the labeling with the help of the labelImg program from python. Along with the training code of the YOLOV7 model, the object detection code was also present. We have modified the basic object detection code so that we can detect the objects and the hand in parallel. On the same code, we added the part for detecting the object closest to the hand. All these changes did not have a significant impact on the speed of hand and object detection. The hand detection model was the one from Tensorflow and there was no need to make changes or train anything custom. But in order to have a higher percentage of success in detecting hands, it was necessary to retexture the hands in the VR environment, they had the default color gray.

Merging steps in a demo

The joining of all components is sequential. I combined the 3d environment with the detection algorithm in the following way: on the Oculus quest 2 vr headset I run the environment as a game. I share the game in real time through the oculus cast feature. The screen being shared is captured by obs and turned into a virtual web camera. This virtual web camera is then provided to the python script related to the detection algorithm. This script does the real-time detection and sets the object to be detected. Depending on the proxy object being tested, the sequence is different from this point on. For the shape changing object these are the following steps. The object to be detected is passed to this script, the script responsible for subordinating the servomotors. This pass is done by simulating a keystroke, because the subscript is waiting for an Input from the keyboard. Depending on the keyboard input, one of the pre-set

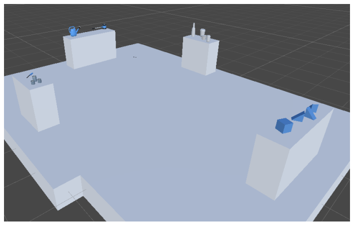

Figure 9: 3D environment for VR.

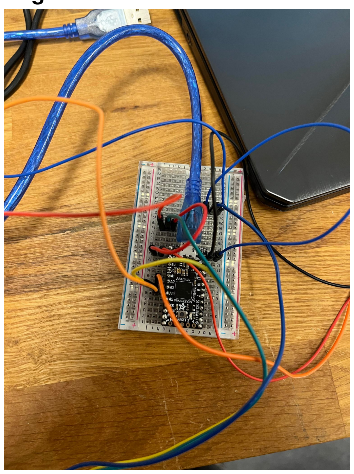

Figure 10: ItsyBitsy m0 connected to servo motors.

wire handling profiles is chosen. For the object consisting of common geometric shapes, the sequence is as follows. Depending on the detected object, it is signaled which region of the object should be extended to the user. Due to time constraints, we didn't implement the robotic arm to do this automatically, so all the related part placement is done manually by one person. This approach can be improved by adding the grab type using GrabNET[5].

Challenges, unexplored possibilities and future direction

During the course of the study, we ran into a number of impediments. One of the problems encountered was setting up the vr headset. Initially we wanted to do the development on an HTC Vive headset, but after a few hours of set-up, we couldn't resist using it, so we switched to the Oculus Quest 2 headset. With it we also had small problems, such as be that he already had a developer account and we had to do a factory reset to be able to use it and we had to create a new developer account. Another problem encountered was the malfunction of the 3D printers. Most of the times the prints failed either because the nozzle was blocked or spaghetti print came out. So it was necessary to learn to adjust the bedings in order to succeed in printing the prototypes, and even after the adjustment, the prints were not guaranteed to be successful

In the 3D environment we encountered a rather significant problem, more precisely the fact that grabbing objects is done using the pinching method, and this is not exactly the most natural way to grab any object, besides that each object has only one the anchor where it can be grabbed, and this means that if the user tries to grab the object from any part without the anchor, the closest part where the anchor is will appear in his hand and this will mislead the user. To solve this problem, enough anchors can be added to the object to cover each surface, but this is a time-consuming solution and increases in complexity with the complexity of the object's shape. Another problem encountered in the vr environment was the simultaneous implementation of hands and controllers. Unfortunately, we gave up on the controllers because we couldn't make them work simultaneously.

Another impediment was the relatively short amount of time, so we didn't manage to develop everything we set out to do and we had to modify some designs and change some ideas. Also for this reason, we were unable to configure the robotic arm and use it in the demo. However, the fact that we didn't use the robotic arm isn't too much of a problem, since it was just a tool to facilitate user interaction.

Related to the unexplored possibilities and future directions, a first direction of development would be the possibility of creating a more advanced grasping system in the VR environment, which would allow a more natural grasping of objects. Another development idea would be to use the grab type to detect the most likely part of an object to be grabbed. In addition to this, work can also be done on improving the method of handling the threads from the object that changes its shape. Also, the object detection algorithm can be improved taking into account the depth in the VR environment.

Evaluation

In this part we test the manufactured objects and check if the designs of the made objects produce the desired result.

Method

In order to evaluate the manufactured objects, we chose to conduct 2 tests, spatial efficiency testing and usability testing.

Spatial efficiency testing

We conducted this test, because it is necessary for manufactured objects to reduce the required storage space. We measured both objects and calculated the volume needed to store them.

Usability testing

This test has the purpose of showing us if the created objects produce the desired sensations or not and if the user can distinguish the figure given to him. This test is conducted in the vr environment and all system components are put into operation for the demo.

Results

Spatial efficiency testing

All the volumes occupied by the objects are in the table 1. After this test, we can say that the least space is occupied by the shapeshifting object, if we do not take into account the structure.

Table 1: Ocupied volum of each object in cm^3

Common shape object	3549 cm^3
Shapeshifting object	36 cm^3
Shapeshifting object with structure	24990 cm^3

Usability testing

We ran this test on a limited number of people, 3 to be exact, because we didn't have time to gather more people. After the test, the opinions were divided and the result is inconclusive, some participants said that for the cylindrical shapes, the common feature object gave a stronger sense of immersion, and for the smaller cylindrical shapes or hammer, the shape shifting object gave a feeling of greater immersion.

Conclusion

In conclusion, although it is not a conclusive result which object is better, the fact that they can produce a greater sense of immersion is a plus and so we can say that we have reached the desired result, an object that improves immersion in the VR environment . Besides that, we also created some objects that occupy relatively little space, between 36 cm^3 and 24990 cm^3

REFERENCES

[1] Bruno Araujo, Ricardo Jota, Varun Perumal, Jia Xian Yao, Karan Singh, and Daniel Wigdor. 2016. Snake Charmer: Physically Enabling Virtual Objects. In *Proceedings of the TEI '16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '16)*. Association for Computing Machinery, New York, NY, USA, 218–226. DOI: http://dx.doi.org/10.1145/2839462.2839484

[2] Thomas Feix, Javier Romero, Heinz-Bodo Schmiedmayer, Aaron M. Dollar, and Danica Kragic. 2016. The GRASP Taxonomy of Human Grasp Types. *IEEE Transactions on Human-Machine Systems* 46, 1 (2016), 66–77. DOI: http://dx.doi.org/10.1109/THMS.2015.2470657

[3] Anuruddha Hettiarachchi and Daniel Wigdor. 2016. Annexing Reality: Enabling Opportunistic Use of Everyday Objects as Tangible Proxies in Augmented Reality. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. Association for Computing Machinery, New York, NY, USA, 1957–1967. DOI: http://dx.doi.org/10.1145/2858036.2858134

[4] Michael L. Rivera, Melissa Moukperian, Daniel Ashbrook, Jennifer Mankoff, and Scott E. Hudson. 2017. Stretching the Bounds of 3D Printing with Embedded Textiles. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 497–508. DOI: http://dx.doi.org/10.1145/3025453.3025460

[5] Omid Taheri, Nima Ghorbani, Michael J. Black, and Dimitrios Tzionas. 2020. GRAB: A Dataset of Whole-Body Human Grasping of Objects. In *Computer Vision – ECCV 2020*, Andrea Vedaldi, Horst Bischof, Thomas Brox, and Jan-Michael Frahm (Eds.). Springer International Publishing, Cham, 581–600.