Interaction Techniques for CSG Modeling in VR

Ricardo Ferreira
ricardo.j.ferreira@tecnico.ulisboa.pt

Instituto Superior Técnico, Lisboa, Portugal

October 2016

Abstract

Many fields use 3D virtual objects and environments which can be modeled using different techniques: sketching, primitives, mesh manipulation, or using CSG, a powerful tool that allows the generation of more complex models using boolean operations on simpler ones. This content is created mainly with traditional 2D input and output. This difference in degrees of freedom to everyday physical interactions is counterproductive. Advances in VR hardware made user tracking and HMD more accurate, comfortable and affordable. VR lets users operate with the same degrees of freedom as they would in the physical world, which has the power to accelerate 3D manipulation and modelling tasks. Moreover, in VR, users can't see their physical body, making it necessary to find strategies to give users the notion of their body, or at least of the parts of their bodies that matter for the tasks they are developing. We proposed two techniques to perform boolean operations between two objects in Virtual Reality based on 3D gestures and menus. To compare those approaches with a baseline approach, with physical controllers, we developed a prototype that implemented all three alternatives and had two types of body representation: a full body avatar or two oval objects for hands. We conducted a user evaluation and results showed that there was no significant difference in efficiency between the three techniques. Users preferred the menu based approach due to its immediate feedback. We also found that having a full body avatar had negatively impacted the efficiency on this type of tasks.
Keywords: CSG, Boolean operations, 3D modeling, Virtual Reality, Self-avatar

1. Introduction

Digital content is used in every field to represent an object or an environment and create new worlds. Virtual objects and environments can be modeled using different techniques: sketching, using primitives, creating the mesh by defining each vertex one by one, or by combining other objects. Boolean operations allow users to generate new and more complex models by combining two simpler ones, a powerful tool to accomplish more complex models faster.

Immersive virtual environments could potentially improve the experience of 3D object modeling and environment creation but we don't have a standard technique to perform those kinds of work in these environments. There are some works related to object manipulation in mid-air, some of them presented in this document, but there are less studies related to object modeling and interfaces to perform those operations in immersive virtual environments, especially to perform boolean operations in mid-air. Furthermore, existing approaches don't make use of gesture tracking for more natural ways of interaction. Instead they still resort to traditional interfaces like 2D menus and physical buttons.

For those reasons, our goal with this work is to study new techniques for modeling 3D objects in immersive virtual environments, with focus on constructive solid geometry, using gesture interactions instead of traditional interfaces while taking advantage of the hardware advancements in head mounted displays and sensors.

While in Virtual Reality, users can't see their own physical bodies because the headsets block their vision. Therefore there is the need to give users a good virtual representation of their position, being that through full body representation with avatars or by representing the parts of the body user needs to perform the task. With our work, we intend to address the following research questions:

1. Are techniques based on natural mid-air gestures more efficient and appealing to users than approaches based on menu or buttons for performing CSG operations in VR?

2. Can self-avatar help users to be more efficient when performing 3D object modeling tasks?

We developed two new ways of performing CSG operations in mid-air for Virtual Reality. The first one is based on the natural gestures that people use to interact with real objects, and the second one is based on traditional menus with the addition of a real time feedback of the boolean operations that are available to choose. To validate those techniques, we performed a comparative test with users where we confronted the techniques with each other and with a third technique,that uses controllers as input, with a button mapped for each of the different operations.

Half of the users performed the test with a full body avatar that mimics all their movements, the rest of the participants carried out the tests with only a virtual representation of their hands. Both portrayals give feedback on position, orientation and gesture of the hands, but only the avatar gives feedback about full body position.

2. Related Work

In this section we are going to evaluate and compare some research related to our work. We are going to classify them their ability of modeling 3D objects, how they do it and finally, by the ability of performing CSG operations and how they do them. Because it is not the focus of our work we will discuss briefly object manipulation. In the end we will also examine the embodiment works that we studied.

We can divide the works in five types of modeling: using primitives [2, 7, 5, 17, 1, 10, 8], sketching [18, 9, 14, 6, 5, 16, 4, 3], extrusion [2, 4, 16, 3], vertex manipulation [2, 16] and using CSG operations[18, 9, 14, 6, 10]. Using primitives is limitative because users are restricted to the available basic shapes to make new objects. With sketching users draw the shapes of the new objects but the system has a limited number of shapes that it recognizes from drawings. Inflation is a very useful technique but the result is generated by algorithms and might not be exactly which artists wanted, requiring more manipulation. Although vertex manipulation allows users to define their objects with greater precision, today there are 3D objects with thousands of vertices and would be very time consuming to place them one by one. Boolean operations are a powerful tool to create more complex objects from simpler ones, like primitives. Works like Sketch [18] or Shapeshop [14] had the ability of taking parts of objects, performing a subtraction operation, but those operations were applied using a single object. MakeVR [10] is the only work, from the ones we studied, that was able to perform boolean operations between two existing objects. This manipulation between two objects offers all four CSG operations (union, both subtractions and interception) and gives a more visual feedback of the result because you can clearly see the volumes of both objects you are manipulating. Because we believe CSG operations are very powerful in 3D object modeling and because we didn't find many techniques to perform them between two objects in mid-air, we decided that this would be the focus of our work, techniques to perform boolean operations in immersive virtual environments.

Manipulating objects is an essential task to model 3D objects and environments. We studied different ways of interacting with the environments. Because finding the better way to manipulate objects in mid-air is a big challenge that is not the focus of our work, we chose to use a well known technique, the 6-Dof Hand [13] which mimics the natural way humans interact with physical objects in the real world. This similarity could potentially make the adaptation of users to the virtual environments easier.

The works studied regarding embodiment showed us that self-avatars change behaviours and the way users interact with virtual environments. In Steed et al. [15] we can see that some users moved their real body to prevent the virtual avatar from beeing hurt by the falling of a box. Lugrin et al. [12] found that realism can have a negative effect in some task performance and that non realistic avatars can achieve the sense of embodiment. Kilteni et al. [11] showed that the avatar look and clothing style can also have an influence in the way users behave in virtual reality. Acknowledging this behaviour influence, we want to find out if avatar has an effect on efficiency in modeling tasks and will compare the performance between an avatar and a non realistic representation of the user's hands.

Although there are some research works about modeling virtual 3D objects in Immersive Virtual Environments, it does not exist any technique to perform boolean operations in VR that uses gestural interactions instead of traditional input. In addition, there is no research work that studies the impact of an avatar in the performance of modeling tasks in Virtual Reality. Therefore, this dissertation's goal is to find an answer to those challenges.

3. Interaction Techniques for CSG Modeling in VR

We propose two new techniques to perform boolean operations in immersive virtual environments, the gesture based and the menu based approaches. In this section we present both techniques and also a third one that is going to be used as a baseline later on.

3.1. Gesture based approach

The gesture based approach is based on gestures that people do in real life. There is no way to perform boolean operations to two objects in our world, therefore there is no gesture that we can translate directly to the virtual world. We interact daily with objects and change their position and orientation often. To perform CSG operations we need to interact with objects and we based our interactions in those real world gestures which became natural and trivial to us because we use them so often. Users bring objects together as they would in the real world and they take the parts of the objects that are not needed to achieve the final result in the same manor.

3.1.1 Creating and manipulating objects

Creating objects is done by grabbing objects from a pallet. This pallet is cast when the user opens its non dominant hand and it follows that hand until it is closed again, with the same gesture. The user can manipulate the position and rotation of the object with the drag gesture in a similar manor to the 6-Hand technique [13]. To grab an object users have to close the hand while touching it, and to drag it you move the hand while its closed. The gesture ends when the hand is opened again. While dragging, the object follows the position of user's hand, and in rotates the same way that their arm does. Scale is done by grabbing an object with a hand and then closing the other hand in a free space. With both hands closed, users can move them away from each other to enlarge the object, or move them closely to make the object smaller.

3.1.2 Boolean operations

To execute the Boolean operations the user has to grab the two different objects on which the operations are going to be performed and drag them together until they are intercepting in the desired position and orientation. When the objects are in the wanted position, the decision phase starts. Objects are divided in three parts: left object less right object, interception and right object less left object. The first is held in the left hand, and the last one is held by the right one. Interception stays at the center whichever movement hands make. The position where the user joined the objects is where the final object is going to be constructed. Users have to move away the undesired parts by draging the objects that they are holding. The right operation will be chosen based on the position where objects where released. The chosen operation can be confirmed by releasing the objects, opening the hands.

In the decision phase, the user can preview the result of the operation by dragging the objects out of the starting position, being possible to return them to that position while the hands are closed. To choose the Union operation, both objects held by the hands have to remain at the starting position. To execute any Difference, the object that is going to be deleted has to be moved away, the other object stays at the starting position with the interception removed. The interception is obtained by moving away both hands leaving the interception at the starting position. To confirm the decision the user opens both hands at the desired position. This gestures are represented in the Figure 1.

3.2. Menu based approach

The menu based approach is inspired by traditional 2D menus present in most WIMP interfaces. We bring this concept to the virtual world representing the menu with a 3D object divided in four sections, one for each selectable options. The menu gives visual feedback to the user by showing previews of the operations in each selectable section, corresponding to the matching operation.

3.2.1 Creating and manipulating objects

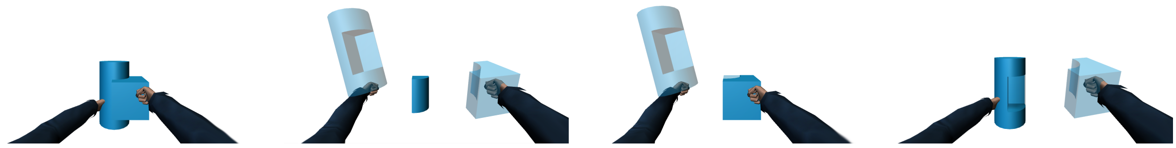
To create objects the user needs to first cast the creation pallet by opening the non-dominant hand. From the pallet the user chooses the wanted object by grabbing it and drags it from the pallet to create a copy. This is done in the same way as in the Gesture approach. Object manipulations in this configuration of the prototype is done in the same way as the Gestures based approach.

3.2.2 Boolean operations

To start boolean operations, the user has to previously have created the two objects between which he wants to operate. Then he grabs one with the dominant hand and drags it until it is intercepting the second object. By releasing the object when it is intercepting another one, the Boolean operations menu is cast and shown on the non-dominant hand of the user, following its movement. In each section there is a preview of the operation that it represents, a little object with the result of the operation on the chosen objects. By touching the menu buttons with the dominant hand , the original objects change to a preview of the result in real time giving the user another way of feedback. By showing this previews the user can choose an operation based on the results instead of choosing the operation itself. To choose which one he desires, the user grabs the correspondent section and the original objects give way to the result of the chosen operation. This process is shown in Figure 2.

3.3. Controller based approach

The controller approach is inspired by the work of Jerald et al. [10], uses handheld controllers with physical buttons and will work as our baseline. In this approach each operation is mapped to a different physical button. The entire arm gesture, positioning and rotation still works like in the other two



(a) Union.

(b) Intersection.

(c) Difference: Left object minus the right object.

(d) Difference: Right object minus the left object.

Figure 1: Possible Boolean operations

(a) Menu appears when releasing an object while colliding with another.

(b) Hovering to preview difference.

(c) Close hand to pick operation.

(d) Operation is applied and menu disappears.

Figure 2: Performing boolean operations with menu based approach.

approaches, but the gestures to grab and drag the objects, to cast of the creation menu and to perform boolean operations are done by pressing buttons on the controller.

3.3.1 Creating and manipulating objects

Objects are created in a similar manner to the other two prototypes but this time around using buttons for input. The creation pallet is toggled on and off with the Menu button. To grab the objects users need to press Grab button, and they have to hold it to drag the object, creating its copy. To hold an object, users press Grab button while touching the object with its digital hand, and hold it while moving the hand to drag.

3.3.2 Boolean operations

To start the operations, users need to grab each object with each hand and bring them together until they are touching each other in the desired position. When the hands stop moving, the objects position is locked and the decision phase begins. Users can now release the objects and move the hands freely. On the controller the user has another four different buttons, one corresponding to each of the Boolean operations. To apply any operations to the objects, the user needs to press the corresponding button on the controller on the decision phase. After pressing, the operation is done and the resulting object appears on the same position of the original objects.

4. Prototype

We built a prototype where we implemented all three techniques explained in the previous Section to test and compare them.

4.1. Architecture

We chose to use the Unity engine because at the time of starting the implementation, it was the most accessible game engine and because all our hardware had official SDK with Unity integration which made it easier to build a first prototype that used all the models from all the hardware that we needed. To calculate the new meshes when performing CSG operations we used a library built for Unity by Andrew Perry. We built a Gesture Manager that decided which gesture the user was performing and did the corresponding action and a User Representation Manager that animated the avatar or hand representation accordingly to user movements.

4.2. Gesture Manager

The Gesture Manager gathers all the information collected through the sensors to decide which is the gesture that the user is performing at each time.

This manager behaviour is implemented in a State design pattern, each state representing each gesture: Idle, Dragging, Scaling, Boolean Operation, Toggle menu visibility and Create Object. In each frame, the sensor data was sent to the current state and it decides if there is the need to change state or not.

The Boolean Operation gesture has slight variations in different approaches but at its base there is the same algorithm, a sequence of three steps: gesture preparation, decision phase, application of the operation. The gesture preparations is where users choose both objects that are going to be affected by this operation by grabbing them with each hand (Gesture based and Controller approaches), or by releasing an object while it touches another (Menu based approach). Decision phase is where you perform the needed steps to choose between the available operations: union, differences or intersection. Application of the operation ends the boolean operations, applies the selected operation on the two objects and generates a new one with the proper result.

4.3. User Representation Manager

There are two possible user representations on our prototype to test the influence of a self-avatar on the performance of our task, one with an avatar, and another without an avatar.

4.4. CSG Module

This module is an implementation of the CSG operations on Unity. Originally this module was built with Javascript by Even Wallace [1] and later ported to Actionscript 3 by Tim Knip [2]. This version is a direct port from the AS3 version to C#/Unity and it was done by Andrew Perry [3].

4.5. Setup

Our setup tracks entire body movement and hand gestures while giving an immersive experience and is present in Figure 3. Using Microsoft's Kinect V2 we were able to get the skeleton of the user to be represented in the virtual world. Although it captures the position for every member, hand gesture and rotation is not achievable with this camera. Instead, we used a Myo armband in each arm to perceive gestures of the hand as well as the orientation of each arm. For the controller version of our prototype we used two Nintendo Wii Remote controllers conected to our PC via Bluetooth. Using a virtual reality headset by Oculus, the Rift DK2, allowed us to give the users a visual immersive experience.

5. Evaluation

In this Chapter we compare the three techniques presented in section 3, and try to find which one is better suited to use as a 3D modeling tool, focusing on Boolean operations between two virtual 3D objects. We asked users to perform a task with each of the different techniques and measured their efficiency and their opinion through an inquiry. We also tested self-avatar influence on user efficiency by executing the tests with half of the subjects with an avatar as a body, and the other half with markers.

[1]https://github.com/evanw/csg.js
[2]https://github.com/timknip/csg.as
[3]https://github.com/omgwtfgames/csg.cs

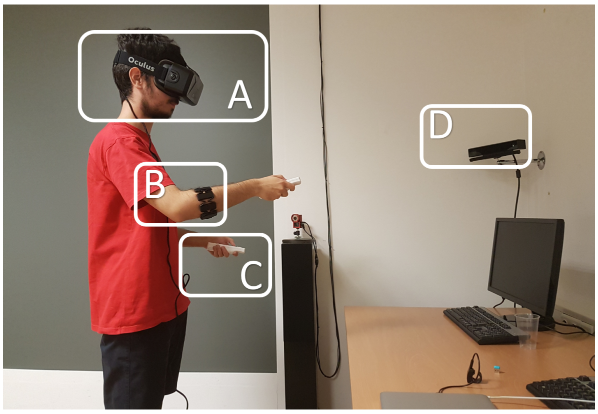


Figure 3: Setup of our prototype. A-Rift DK2; B-Myo armband; C-Wii remote controller; D-Kinect v2.

5.1. Participants and apparatus

The tests were performed in our laboratory in the Taguspark campus of Instituto Superior Tcnico, also known as Loureno Fernandes Laboratory. The room has restricted access which produces an isolated, clean, calm, and controlled environment without external disturbance. This room is fully equipped with all the hardware needed to carry out the tests: Kinect cameras, Myo armbands, Oculus Rift DK2, Wii Remotes and a projection screen for the introductory presentations. We performed the tests with 24 people (3 female), with ages from 21 to 30 years old. Most were students in higher education (75%), while the remainder had already achieved a MSc degree. Less than half (42%) had never experienced Virtual Reality before, only 4 people (17%) had never experienced 3D gesture tracking systems like Microsoft's Kinect, Playstation Move or Wii Remote Controllers, and only 3 people (12.5%) had never experienced modeling tools.

5.2. Methodology

Tests were carried out individually. An example of users performing the test can be seen in Figure 4. For each one of them, we gave a small presentation explaining how the prototype works, a script with instructions and asked them to do a task. Every user performed the task three times, one for each technique: gestures, menu, controller. The order of the techniques they used was changed in every test, going through a total of 6 different possible orders four times. To test self-avatar efficiency, half of the tests were done with an avatar representing the whole body of the user on the virtual world, and the rest were done with two round objects representing only the hands. In the end, after performing the task with all three techniques the users answered some more global questions, comparing

(a) Participant using the gesture and menu based approaches.

Figure 4: Users performing the task

all three techniques and their version of body representation.

5.3. Task

The task that users were asked to do is to replicate a 3D model with our prototype. The model is represented in Figure 5.

(a) View from the top. (b) View from the bottom.

Figure 5: Objective model

5.4. Results and Discussion

To evaluate the difference between the three techniques to perform CSG operations in Virtual Reality, we collected objective and subjective data in the form of logs and inquiries respectively. To test data normality we used the Saphiro-Wilk test. With the normal distributed data we found if there were significant differences with the ANOVA with repeated measures test with a Greenhouse-Geisser post-hoc test and a Bonferroni correction (corrected sig. = sig. x3), when data wasn't normal distributed we used Friedman non-parametric test with Wilcoxon-Signed Ranks post-hoc test, also with Bonferroni correction. To test the significant differences in embodiment, we used the independent t-test.

5.5. Objective Data

With efficiency evaluation of the three techniques in mind, we collected the times (in seconds) of users performing the task with each one. Our logs kept the total time (Figure 6) of execution of the task as well as individual logs for each kind of operations like object manipulation, Boolean operations, object creation and idle (Figure 7).

We found statistically significant differences between approaches in the total completion time of the task ($F(1.874,37.481)=5.101, p=0.012$). With the pairwise tests comparing total time, we found that there was statistically different between the Gestures and Menu alternatives ($p=0.023$). This significant difference in total time lead us to test the time for individual type of manipulations to find where that difference was more prominent. We found with this that there was no significant difference in the execution of Boolean operations. Differences were felt in manipulation operations($F(1.718,36.076)=6.499$, $p=0.006$) with post-hoc test revealing that there was statistically significant difference between Gestures and Menu approaches ($p=0.062$) and between Gestures and Controller approaches ($p=0.018$). Idle times had also a significant difference ($\chi^2(2)=7.128$, $p<0.05$). We ran Wilcoxon-Signed Ranks to find that the Gesture and Menu approaches are statistically different in Idle time ($p=0.057$).

Although we found no significant difference in execution of boolean operations, we have to recall that the gesture based and menu based approach used the Myo arm bands as input in contrast to the Wii remote controllers used in the controller based technique. Myo arm bands gave us the enormous vantage of having free hands but also brought inaccuracy. Wii remote controllers worked as expected every time, Myo arm band, on the other hand, detected wrong gestures which introduced errors and lead to user mistakes and ultimately to frustration and worst execution times. This observation leads us to think that this interfaces could have advantages when compared to the controller approach if the hand-gesture detection hardware was better. The differences felt manipulation-wise could be caused by users being faster with the controllers because of its higher accuracy, or because users tend to better prepare the position of objects to execute boolean operations when using myo armbands. To better find out, we did some pairwise tests to find where those differences were felt. Idle times differences could also be result of users having to think more in some approaches that others or even by users having to rest more, or being more frustrated by the errors caused by myo arm bands.

By testing approaches pairwise we can see where there are the differences. Total times between gesture based and menu based alternatives are influenced by the difference in manipulation. In spite of the fact that object manipulations gestures are exactly the same in gesture based and menu based, the differences in the boolean operations are enough to influence that significant difference in times. On the gesture based approach, because users had to grab objects with both hands, one on each hand, and because there was a Myo armband in each arm, users tended to make more mistakes. To compensate that, users carefully positioned objects to better prepare boolean operations, resulting in bigger times of manipulations. Also, on menu based approach, it was easier to cancel boolean operations and re-adjust object position, which also made users more careless on manipulation gestures, and therefore faster. The difference between gesture based and menu based approaches could be justified by the efficacy of the Wii remote controllers when compared to the Myo armband gesture detection. That advantage contributes to more efficiency in executing the operations, allowing users to perform them faster and with less preparation and thought, being that a key factor to the difference in Idle time as discussed before.

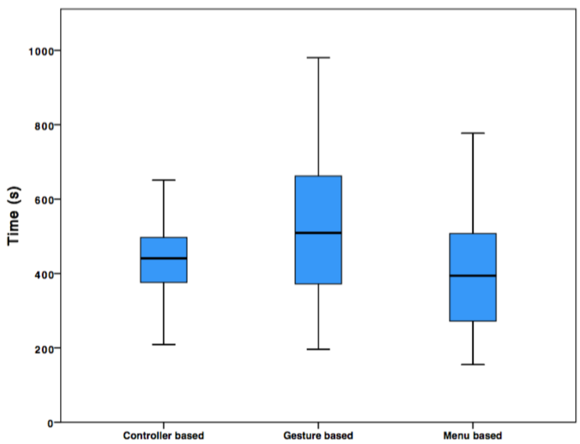


Figure 6: Test total times for the different techniques

Our embodiment test was done with an independent-samples t-test and we found that not only there is a significant difference time between having an avatar or not in total times (t(22)=-3.003, p=0.007), there is also a difference in each approach of the prototype between having an avatar or only hand representation: gestures (t(22)=-1.792, p=0.087), menu (t(22)=-1.72, p=0.099) and controller (t(21)=-2.927, p=0.008).

These results tell us that the self-avatar had a negative impact on the performance of users executing our task. This conclusion disagrees with the studies and theories about embodiment that we referred in section 2. We have some some suggestions to help us understand why we achieved this results.

Avatars are great to help users feel present in a virtual world where they need to see their scale, and to move through the immersive virtual environment. While our work gave the possibility of movement through the room, most users stayed in the same place where they started the test, moving only their arms to interact with the objects. Therefore the advantages of seeing our virtual size and where we are in the world in relationship to the ground were not felt at all.

The only real part of the avatar that was used by every participant was the hands, because they had to use them to interact with the objects. When the avatar was not present, there still was a representation of the hands, by an oval object, which gave the users enough perception of position to operate the objects. As a matter of fact, the change of state of those oval hand representation was more noticeable because they changed colors, than on the avatar counterpart, where the change was only felt by the position of the fingers. Our participants could have perceive the change of color faster than the change of position of the fingers. So, although there was no avatar, there was still the same amount of feedback present.

Avatar arms potentially occluded the vision of users by being in front of objects blocking them and the world, affecting overall perception of the scene. Lastly, avatars can have raised wrong expectations in users. By having arms and hands similar to theirs in the real world, they create the idea that they move the same one that real ones do. Virtual movements were constructed based on the capture of a Kinect v2.0 camera which introduces some imprecision and jitter. This might cause some frustration in users that wanted to move the arms like real ones. Participants with ovals only had approximations of the hand, and did not expected them to move like real hands because they didn't look familiar, adapting faster to the jittering movement.

5.6. Subjective Data

We collected subjective data to help us identify the user's opinion, their preferences and other useful information that is not collectible through automatic logs. To compare the different approaches we asked the users how difficult and fun was to perform the task with each of the prototypes. To find how different was to execute the operations we asked them how hard was to recall how to do them and how hard was to actually perform them. Those preferences are represented in Table 1.

To test if there was a significant difference we used Friedman non-parametric test with Wilcoxon-Signed Ranks post-hoc test and with Bonferroni

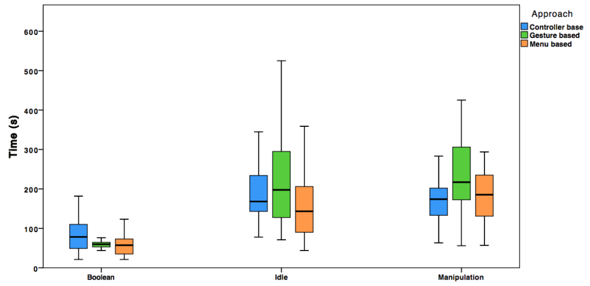


Figure 7: Pairwise comparison of times for the different techniques separated by operation.

correction. We found statistically difference in every parameter: difficulty , fun, recall and execution.

Although these results represent the overall opinion of participants, we do think that their judgement was heavily influenced by the challenges that the low efficacy of the input introduced. Kinect input is an approximation and because of that it has some very noticeable jumps from frame to frame. Myo armbands sometimes failed to proper recognize the right hand gestures that testers were performing even though they were calibrated for each one of them. This jitter and imprecise controls lead to frustration that was more felt in one particular approach than others.

In terms of difficulty, we found significant difference (χ^2 (2)=7.786, p <0.05) between approaches and with the pairwise tests that there was statistically difference between the gesture based approach and the menu based approach(p = 0.012). This difference was felt because the menu approach gives more immediate feedback. As soon as the boolean menu shows up, users are presented with the results of every boolean operation, making more direct which one they have to choose when compared to the thinking or experimenting that they have to do in the gesture based approach. Also, because they didn't have to hold the objects in the menu approach, they could re-adjust the position of objects instead of choosing an operation. This difference could also be influenced by the performance of the input hardware. To perform boolean operations in the gestures approach, it was necessary to hold the objects with both hands and for more time than on the menu approach. Myo armbands recognized, often, hands opening when they were in fact still closed which had a big impact on the final result of the operation. The preview was done while holding the objects and because users had fear that the armbands would wrongly detect that they opened their hands, they wouldn't depend on that preview to execute the operation. They often preferred to think before in what they have to do. On the menu based counterpart of the prototype, users only had to hold objects with one hand, therefore less possibility for hardware failure, and preview is done after positioning object, without holding objects which gave testers more confidence to play with. Also, there is the possibility of re-adjust object position in the menu based approach and that fact makes the Myo wrong detections less significant because they could be corrected when it happened. That adjusting component also compensates the jitter introduced by Kinect which is more felt on gesture based approach because it lacks that possibility of readjustment.

Fun factor also had significant difference (χ^2 (2)=6.123, p <0.05) with pairwise tests reveiling differences between gesture based and menu based approaches (p = 0.42). All those problems introduced by the hardware were very frustrating to the participants when they were performing the task with the gesture based approach. The jitter when positioning the objects and the lack of re-adjusting their position made the errors more impactful on the final result and that had a very negative impression on users enjoyment. For these reasons, the gesture based approach was less fun that the menu based for most of participants.

There were significant differences for Recall(χ^2 (2)=30.104, p <0.05) and also differences between every approach: Gestures and Controller (p = 0.006), Gestures and Menu (p = 0.021) and Controller and Menu (p = 0). Recall is the only parameter were we can compare directly all three variations of the boolean operations, because it has the only results where we found statistically difference between all approaches in pairwise tests. Menu based is the easier to recall not only because it showed the result of each operation in the correspondent button, but it also showed the final result on the actual objects by hovering the buttons with the dominant hand. Users could choose the boolean operation needed to achieve the wanted result without really knowing which operation they were choosing. Gesture based were the second easiest because they also showed the preview of the result of the operation that users are choosing by changing the material of the parts that were going to be deleted to a translucent blue. With the low efficacy of the hardware used, participants were not as comfortable using this preview as they were in the menu based approach. This also helped to the fact on making menu easier than gestures. Controller is the last approach in this recall ranking because it doesn't possesses any kind of preview feature. Participants had to remember which physical key they had to press corresponding to each of the operations. There was an image on the virtual world that they could consult in case they forgot.

We found statistically significant difference in

	Gesture based	Menu based	Controller based
Easiness*	4.5 (2)	5 (1)	4.5 (1.75)
Fun*	5 (1)	6 (1)	5 (1.5)
Recall*	5 (2)	6 (1)	3 (2.5)
Execution*	4 (2.75)	5.5 (1)	4.5 (2.75)

Table 1: Participants opinion for each technique (Median, Inter-quartile range). * means that there were statistically significant differences for that factor.

terms of execution (χ^2 (2)=11.541, p <0.05) and the Wilcoxon-Signed Ranks post-hoc test showed us significant between Gestures and Menu for (p = 0.03). With our results we can say that users found executing the boolean operations easier in the menu based approach than on the gesture based approach. This is consequence of the stronger feedback menu gives to the users with two types of previews. Before choosing an operation, users have a preview of all operations on the sections of the menu, and there is also the possibility of hovering those sections to have a real time preview on the actual objects. Users had to think and experiment less with menu based approach than with the gesture based approach and they could also re-adjust the position of objects before making a selection. That difference was also influenced by the hardware as discussed before. The frustration felt by the users made them fear the preview feature on the gesture based approach which made it more difficult than we thought it would be, in theory, if the input hardware was more reliable. Also, on the gesture based technique, even though participants didn't experiment as much with the four different possibilities, they would still have to perform the gesture of that particular operation which was prone to input errors making the task more difficult.

When asked which technique they liked most and would like to see in a CAD like application in VR, 50% of people answered the menu-based approach, 37.5% chose gesture-based approach and the left 12.5% chose the controller based approach.

6. Conclusions

We designed two techniques to perform boolean operations between two 3D objects in immersive virtual environments. The first technique is based on gestures humans do to move objects in the real world. The second technique is based on traditional menus with buttons. We also had a third technique that we used as a baseline. This technique uses handheld controllers with physical buttons as input.

The results of user tests have shown us that are no differences in efficiency between boolean operation performance with the three techniques. The Menu approach was easier to use, easier to recall and to perform boolean operations and more fun to use than the gesture based technique. With controller based approach it was more difficult for users to remember how to perform boolean operations than with the other two techniques. Most users found that the better technique was the menu based approach.

Our prototype had two modes to represent the body of participants. We gave each representation to half of the participants and compared the results from each group. We found that self-avatar had a negative impact in the total time users took to perform the task with every approach.

6.1. Future Work

In our work there are some aspects that can be improved in future or be the focus of future works:

Myo tracking is not perfect and introduces some false positives and also some false negatives, which introduces errors and increases times in task performance, and tracks rotations of the arm instead of the hands. We believe that for object manipulation, tracking the rotation of the hand is more appropriate.To use the gesture based technique with this type of tracking we suggest an alteration to the technique so it doesn't require users to grab the objects for too long.

By evaluating the quality of the models created by users, it is evident that they are far from perfect. They are not properly aligned and the proportions are not ideal in most cases. To get better results we suggest combining our techniques with object manipulation approaches that enable users to get more precise object placement and rotations.

Controller based approach had the disadvantage of having no preview of the boolean operations result. This can be changed in future work by implementing previews in this technique.

References

[1] C. Barot, K. Carpentier, M. Collet, A. Cuella-Martin, V. Lanquepin, M. Muller, E. Pasquier, L. Picavet, A. Van Ceulen, and K. Wagrez. The wonderland builder: Using storytelling to guide dream-like interaction. In *3D User Interfaces (3DUI), 2013 IEEE Symposium on*, pages 201–202. IEEE, 2013.

[2] J. Butterworth, A. Davidson, S. Hench, and M. T. Olano. 3dm: A three dimensional modeler using a head-mounted display. In *Proceedings of the 1992 symposium on Interactive 3D graphics*, pages 135–138. ACM, 1992.

[3] D. Cochard, P. Rahier, S. Saigo, and M. F. Maltouf. Building worlds with strokes. In *3D*

User Interfaces (3DUI), 2013 IEEE Symposium on, pages 203–204. IEEE, 2013.

[4] B. R. De Araùjo, G. Casiez, and J. A. Jorge. Mockup builder: direct 3d modeling on and above the surface in a continuous interaction space. In *Proceedings of Graphics Interface 2012*, pages 173–180. Canadian Information Processing Society, 2012.

[5] M. F. Deering. The holosketch vr sketching system. *Communications of the ACM*, 39(5):54–61, 1996.

[6] A. S. Forsberg, J. J. LaViola Jr, and R. C. Zeleznik. Ergodesk: a framework for two- and three-dimensional interaction at the activedesk. In *Proceedings of the Second International Immersive Projection Technology Workshop*, pages 11–12. Citeseer, 1998.

[7] T. Ha and W. Woo. Arwand for an augmuented world builder. In *3D User Interfaces (3DUI), 2013 IEEE Symposium on*, pages 207–208. IEEE, 2013.

[8] K. Hald. Low-cost 3dui using hand tracking and controllers. In *3D User Interfaces (3DUI), 2013 IEEE Symposium on*, pages 205–206. IEEE, 2013.

[9] T. Igarashi, S. Matsuoka, and H. Tanaka. Teddy: a sketching interface for 3d freeform design. In *Acm siggraph 2007 courses*, page 21. ACM, 2007.

[10] J. Jerald, P. Mlyniec, A. Yoganandan, A. Rubin, D. Paullus, and S. Solotko. Makevr: A 3d world-building interface. In *3D User Interfaces (3DUI), 2013 IEEE Symposium on*, pages 197–198. IEEE, 2013.

[11] K. Kilteni, I. Bergstrom, and M. Slater. Drumming in immersive virtual reality: the body shapes the way we play. *IEEE transactions on visualization and computer graphics*, 19(4):597–605, 2013.

[12] J.-L. Lugrin, J. Latt, and M. E. Latoschik. Avatar anthropomorphism and illusion of body ownership in vr. In *2015 IEEE Virtual Reality (VR)*, pages 229–230. IEEE, 2015.

[13] 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.

[14] R. Schmidt, B. Wyvill, M. C. Sousa, and J. A. Jorge. Shapeshop: Sketch-based solid modeling with blobtrees. In *ACM SIGGRAPH 2007 courses*, page 43. ACM, 2007.

[15] A. Steed, S. Frlston, M. M. Lopez, J. Drummond, Y. Pan, and D. Swapp. An in the wild-experiment on presence and embodiment using consumer virtual reality equipment. *IEEE transactions on visualization and computer graphics*, 22(4):1406–1414, 2016.

[16] T. M. Takala, M. Mäkäräinen, and P. Hämäläinen. Immersive 3d modeling with blender and off-the-shelf hardware. In *3D User Interfaces (3DUI), 2013 IEEE Symposium on*, pages 191–192. IEEE, 2013.

[17] J. Wang, O. Leach, and R. W. Lindeman. Diy world builder: An immersive level-editing system. In *3D User Interfaces (3DUI), 2013 IEEE Symposium on*, pages 195–196. IEEE, 2013.

[18] R. C. Zeleznik, K. P. Herndon, and J. F. Hughes. Sketch: an interface for sketching 3d scenes. In *ACM SIGGRAPH 2007 courses*, page 19. ACM, 2007.