Consiglio Nazionale delle Ricerche
Istituto di Matematica Applicata
e Tecnologie Informatiche
"Enrico Magenes"

REPORT SERIES

E. Cordeiro, F. Giannini, M. Monti, A. Ferreira

A study on 3D shape interaction in virtual reality

18-02

IMATI REPORT Series

Nr. 18-02
April 2018

Follow this and additional works at: http://www.imati.cnr.it/reports

IMATI Report Series Nr. 18-02
April 3, 2018

A study on 3D shape interaction in virtual reality

Eduardo Cordeiro, Franca Giannini, Marina Monti, Alfredo Ferreira

Eduardo Cordeiro
Istituto di Matematica Applicata e Tecnologie Informatiche "*E. Magenes*"
Consiglio Nazionale delle Ricerche
Via de Marini, 6 (Torre di Francia) 16149 Genova - Italy
E-mail: eduardo.cordeiro@ge.imati.cnr.it

Franca Giannini
Istituto di Matematica Applicata e Tecnologie Informatiche "*E. Magenes*"
Consiglio Nazionale delle Ricerche
Via de Marini, 6 (Torre di Francia) 16149 Genova - Italy
E-mail: franca.giannini@ge.imati.cnr.it

Marina Monti
Istituto di Matematica Applicata e Tecnologie Informatiche "*E. Magenes*"
Consiglio Nazionale delle Ricerche
Via de Marini, 6 (Torre di Francia) 16149 Genova - Italy
E-mail: marina.monti@ge.imati.cnr.it

Alfredo Ferreira
Instituto Superior Técnico,
Dept. Eng. Informática
Av. Rovisco Pais, 1 - 1049-001 Lisboa -Portugal
E-mail: alfredo.ferreira@tecnico.ulisboa.pt

Abstract.

Current immersive modelling environments use non-natural interfaces to support traditional shape modification operations. In the future, we expect that natural interfaces will become the typical interaction method for 3D shape modelling. In this report, we present a study on user gestures and speech interaction for shape modification. With this experiment, we aim to study how users model and manipulate three-dimensional objects in a virtual reality environment through gestures and speech without using any additional device, i.e. no pointer or controller. It will give us a detailed and accurate perception of how users can interact naturally with three-dimensional objects in a virtual reality environment.

Keywords: *Natural interfaces for shape modelling; Immersive environment; mid-air gestures*

Sommario

The Objective of the study 3
What will be evaluated 3
Environment 3
Tasks to be performed 4
Methodology 9
Wizard of Oz Testing 9
Hardware 9
Participants 10
Subjective Analysis 10
Experiment Results 10
User Strategy 11
Results and Discussions 11
APPENDIX 13
Appendix A - Procedure to be followed for the experiment 13
Appendix B - Video Analysis 15
APPENDIX C – Profiling Questionnaire 21

The Objective of the study

The objective of the study is to understand how users model and manipulate three-dimensional objects in a virtual reality environment through gestures and speech without using any additional device, i.e. no pointer or controller.

What will be evaluated

Object Manipulation and Object Modelling

Our objective is to identify a set of gestures and voice commands performed by users during manipulation and modelling tasks.

Our goal is to gather information in regards to the types of gestures and/or voice commands used during the test, in order to develop a prototype for immersive object modelling. We will perform an objective and subjective study on the manipulation and modelling of 3D objects in immersive environment.

We intend to determine what types of gestures and voice commands users employ when performing these tasks.

Environment

The users are able to perform any gesture and voice command they want, as there are no restrictions imposed. We want to understand how users perform the tasks in the most natural manner.

The tests will be conducted in the multimedia laboratory of IMATI-CNR, Genova, Italy. The room is equipped with all the necessary peripherals that are required for the test: VIVE headset and cameras and a Leap Motion.

The room has restricted access, so the environment where tests will be performed is controlled and without any external disturbances

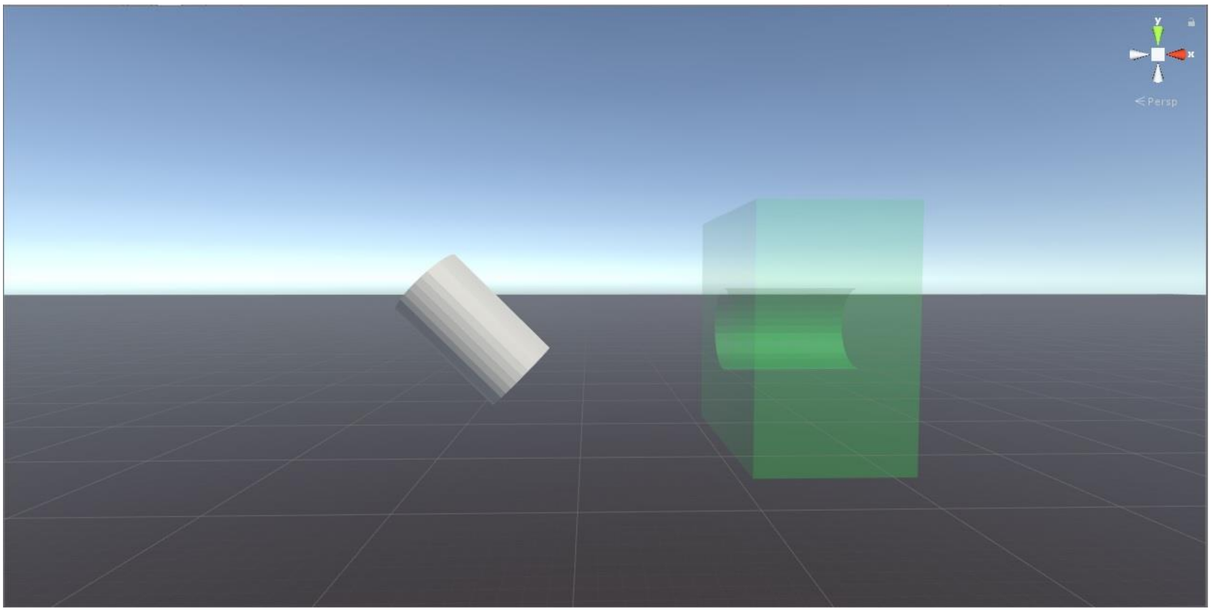
Participants will be asked to complete a set of tasks, which consists of five different modifications to an object in the virtual environment: we chose tasks based on previous work Cui et al. (2016b)[1], as they represent a subset of the object modification tasks presented.

[1] Cui J, Fellner DW, Kuijper A, Sourin A. Mid-air gestures for virtual modeling with leap motion. In: International Conference on Distributed, Ambient, and Pervasive Interactions. Springer; 2016a. p. 221{30.

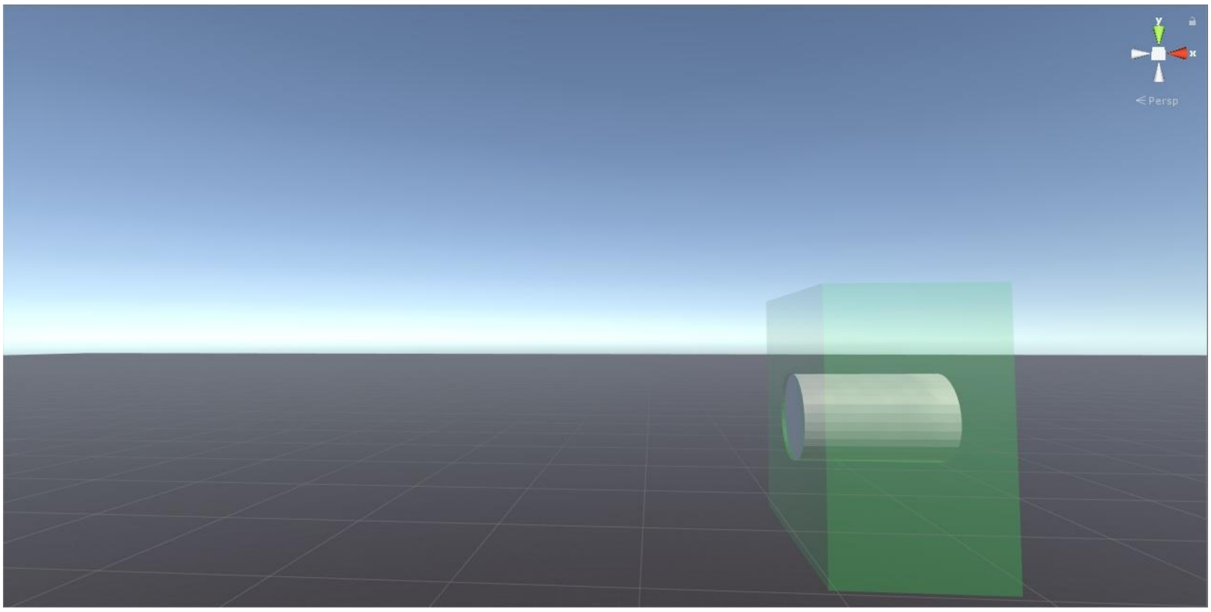
Tasks to be performed

Task 1 (Docking):

- This tasks requires the user to move and rotate a Cylinder (gray), using gestures or voice commands, and place it in the appropriate position inside the box (green)
- Start of the Task:

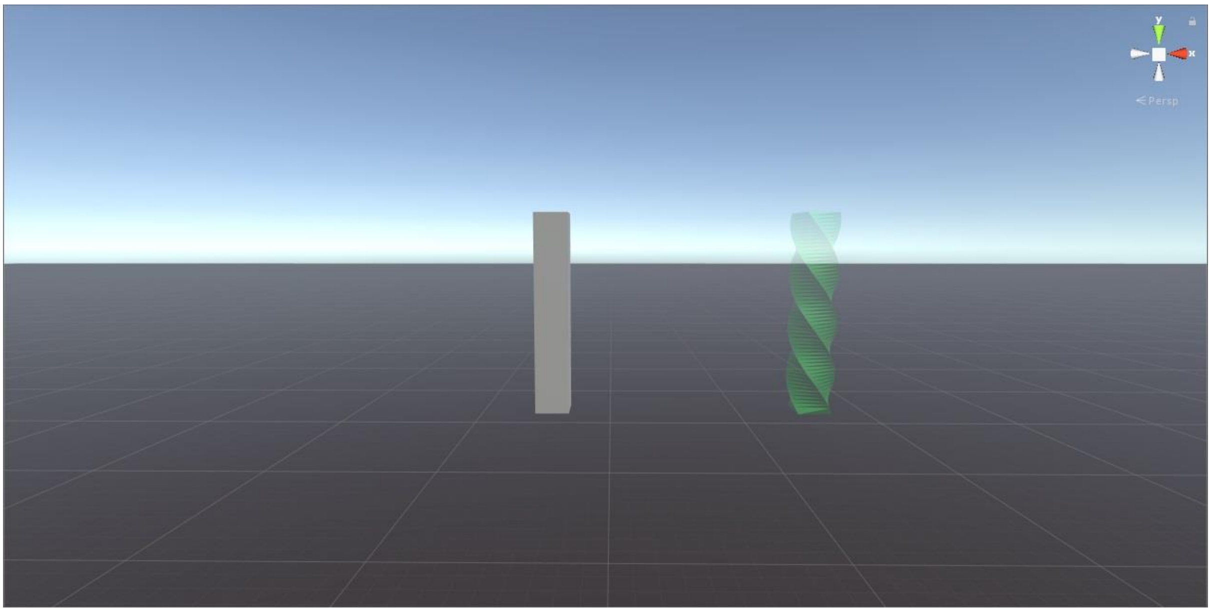


- After Completion:

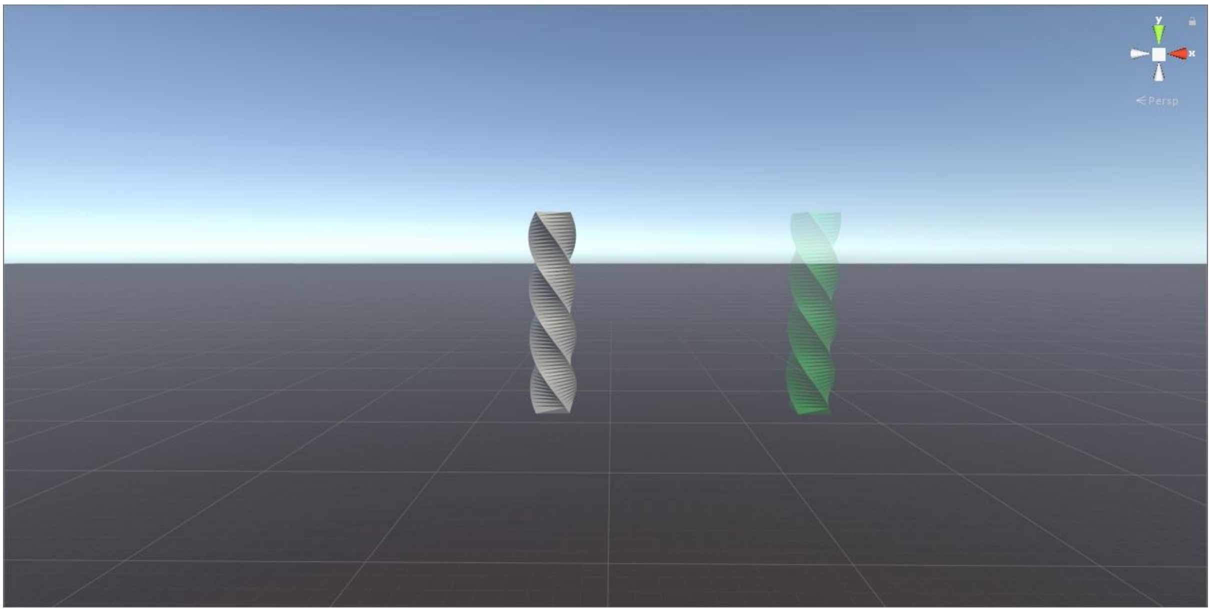


Task 2 (Twisting):

- The user has to perform gestures or voice commands in order to make the starting object (parallelepiped, in gray) into the twisted object (green)
- Start of the Task:

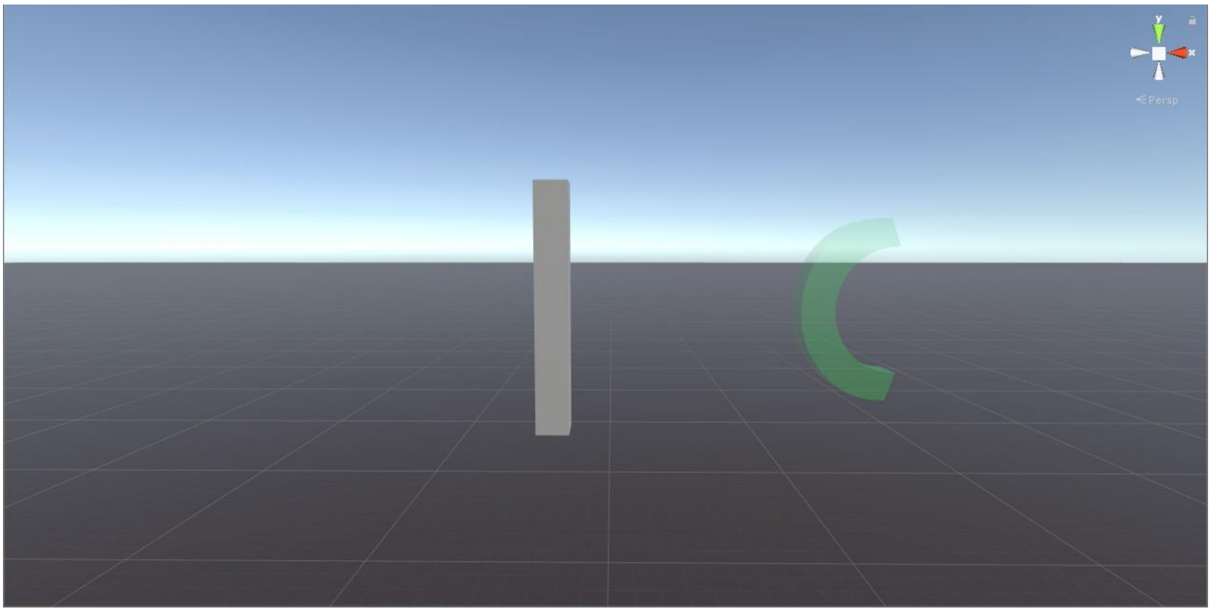


- After Completion:

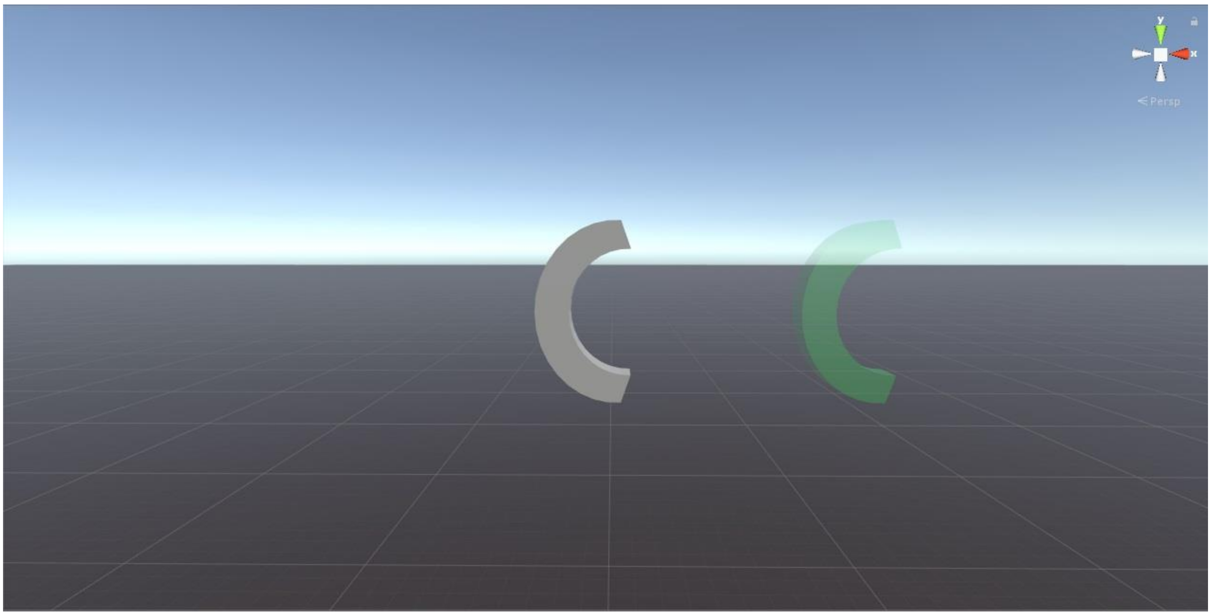


Task 3 (Bending):

- The user has to perform gestures or voice commands in order to make the starting object (parallelepiped, in gray) into the bent object (green)
- Start of the Task:

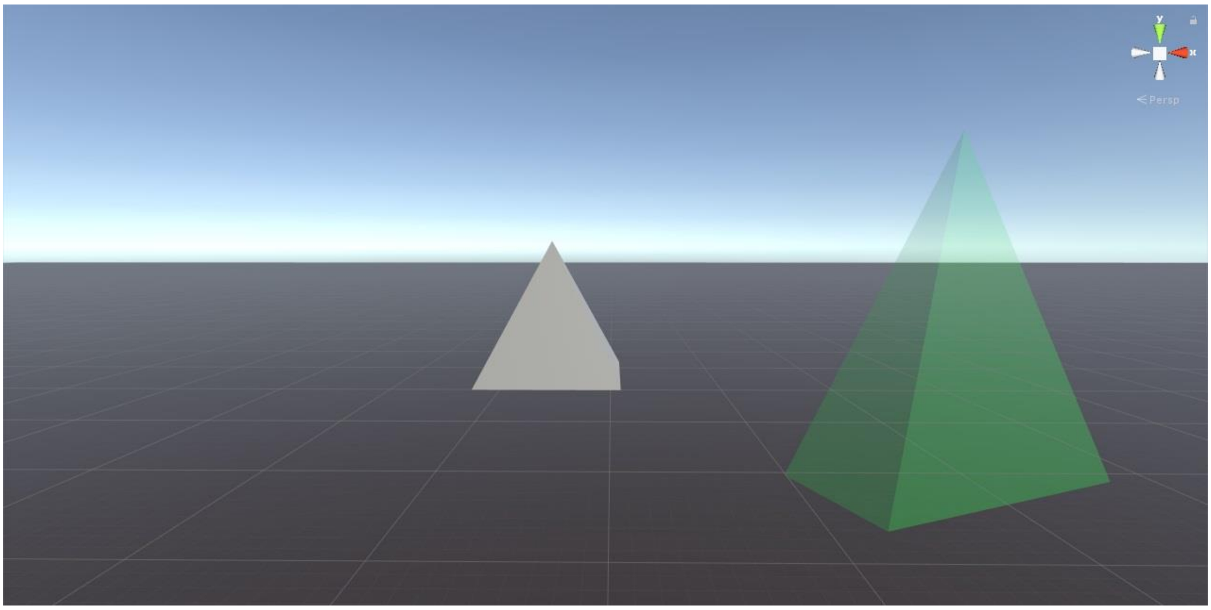


- After Completion:

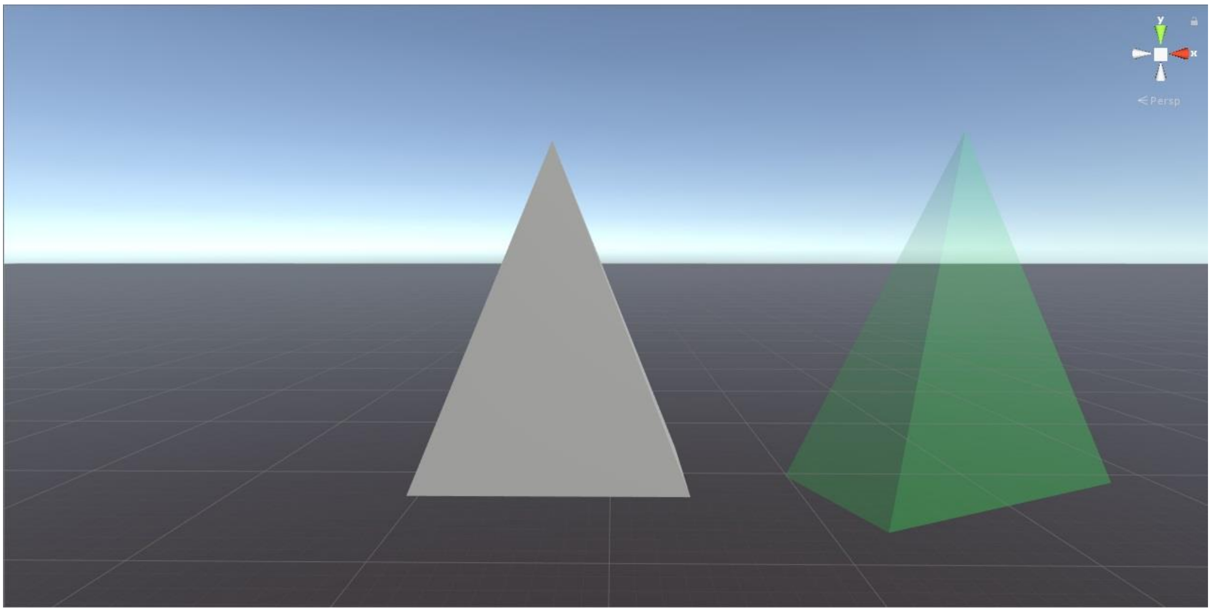


Task 4 (Scaling):

- The user is required to perform gestures or voice commands on the starting object (gray) in order to achieve the desired object (Scaled version, green)
- Scale an object to match the desired object (uniform and non-uniform)
- Start of the Task:

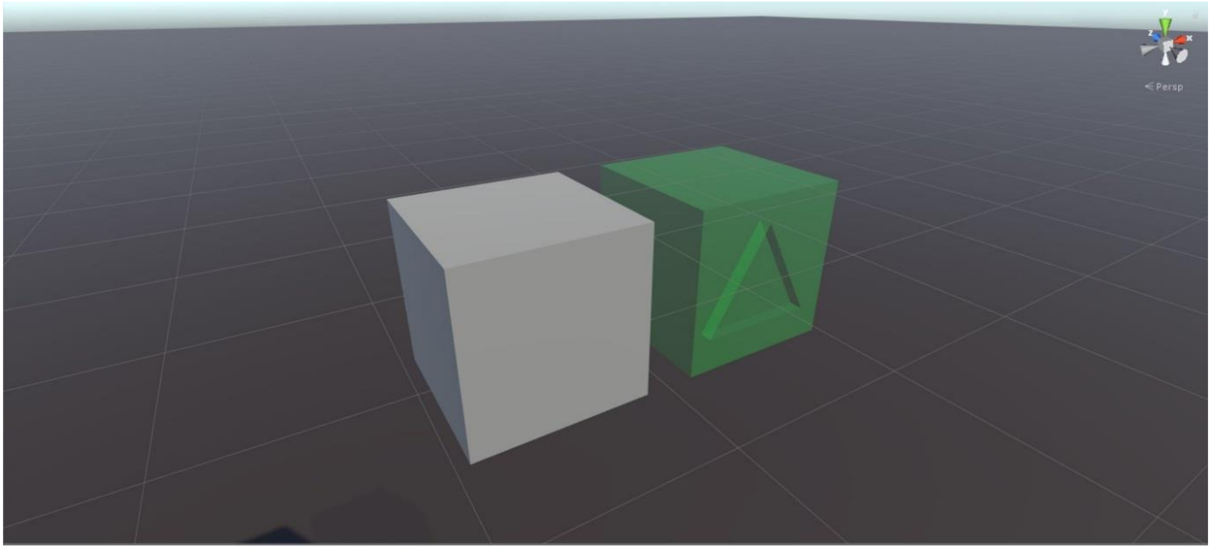


- After Completion:

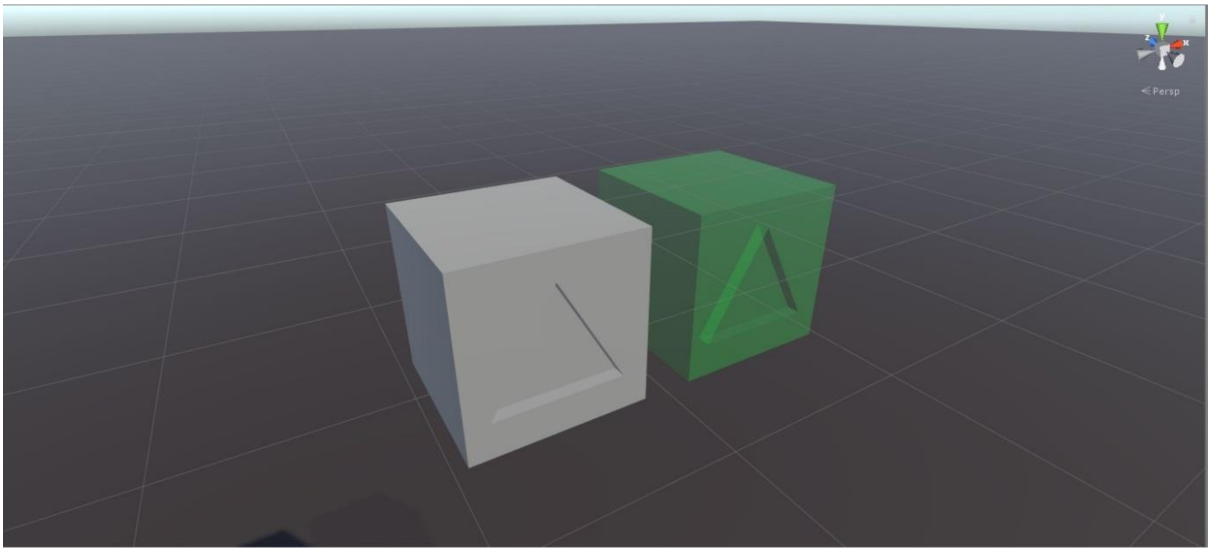


Task 5 (Depression):

- The user is asked to perform gestures or voice commands in the starting object (Gray) that result in the target object (green)
- Start of the Task:



- After Completion:



Methodology

Each user evaluation session followed the same methodology and lasted approximately 24/30 minutes. The experiment began with a short introduction on the aim of the test and what was expected from each participant. Each person was then asked to fill out a profiling questionnaire (appendix C). Afterwards a script was shown with the tasks and the objectives of the experiment that the participant would perform. Participants were given a training period to adjust themselves to the virtual environment and to the virtual representation of their hands. Subsequently, participants were instructed to perform the five tasks described above. The order of the tasks presented to the user followed a partial random order, so that all permutations were exhausted, to avoid biased results.

Appendix A illustrate the schema of the procedure followed for the experiment.

Wizard of Oz Testing

As noted previously, we conducted a Wizard of Oz user experiment. We chose this approach as we were interested in understanding how the human motions are performed in a scenario without almost no constraints. In this experiment the user actions are performed by a human operator from behind the scenes, without the user being aware of this. To achieve this we created a set of objects that corresponded to different steps of the manipulation and modelling of the objects in each task.
During each task, the operator was able to see on a screen what the user was seeing in head-mounted display, interpreted the user actions and simulated the results of the hand motions and voice commands, making the corresponding changes to the object. To simplify the interaction process we chose to have the operator select the objects for the user. Users were told to place one of their hands inside the object for a brief duration (1-2 seconds) to select the object, and we allowed direct manipulation of the object that was attached to the hand. Resting their hands inside an already selected object would yield the inverse result, a de-selection. The wait period gave the operator time to understand if the user wanted to perform a selection or de-selection, without signalling to the user that he was not really doing the action.

Hardware

The experiment was carried out in a controlled environment; the operator is sitting in front of the desktop computer and the participant is wearing the head-mounted display used in the experiment. The fundamental purpose was to correctly gather information from the user hand motions, and to provide a pleasant visualization experience to the participants. To retrieve the pertinent information from the user hand motions we used a Leap Motion sensor. It features two cameras and three infrared LEDs, which are used to track infrared light. For the visualization component an HTC VIVE head-mounted display was used, in combination with two VIVE cameras. This headset is capable of tracking the orientation of the user head using three degrees of freedom. A standard point-and-shoot video camera was used to record the participants for a later analysis. The data recorded allowed us to analyse the hand and voice interaction, in order to obtain meaningful results. Additionally, the operator of the tests used a Desktop PC with a common keyboard and mouse input to perform the necessary actions during the user tests. The prototype used in the user tests was developed using the Unity 3D engine.

Participants

The experiment involved 21 participants, 13 male and 8 females, with ages evenly distributed between 18 and 60. The strong majority held at least a Bachelor degree (90.4%). Most of them with a background in Computer Science or Applied Mathematics, some of which with strong knowledge of shape representation and processing, and some with experience of shape modelling tools. When inquired about their experience with Virtual Reality, 95.2% mentioned that they tried it at most once. This was also the case for Gesture Detection systems (Leap Motion) usage, as 95.3% responded in the same way.

Subjective Analysis

A subjective analysis has been also carried out including:

- Questionnaire to understand what users think of the operations
- Demographic questionnaire
- Informal interview (5~10 mins / user) to gather additional information

Experiment Results

In this section we present the results from the user experiment. The data used to obtain these results was gathered via the video recordings, the post experiment interviews and the profile questionnaire.

Considering that participants often employed more than one interaction method to complete a single task, we decided to separate the respective modes.

For example, if a participant completed one task using both a uni-manual approach and voice, we will attribute a 0.5 value to each of those methods. In the Docking task 100% of participants used an uni-manual approach. For the Twisting task we found that the majority (71.4%) used a symmetrical approach, 19% used only one hand to perform the task and 9.5% employed an asymmetrical style. In the Bending task we found that 57.1% preferred the symmetrical mode, 28.6% chose to use an uni-manual approach and 4.8% adopted an asymmetrical method. Additionally, 2 participants used voice commands to complete this task. We found similar results when analyzing the Scaling task, as 66.7% of participants also chose the symmetrical mode, 38.1% employed an uni-manual mode, 14.3% adopted the symmetrical approach and one participant resorted to voice commands in this task.

In the Local Deformation task, the data show that the great majority (81%) adopted the uni-manual mode, while 14.3% chose to use a symmetrical approach.

We also retrieved data regarding the use of voice interaction during the experiment; For the Docking and Twisting task no participants used voice to express any command or comment. In the Bending task we found two participants that resorted to voice interaction. One chose to only use voice to complete the task, while the other performed a gesture in combination with a description of what she was performing. In the Scaling task, only a single participant used voice to achieve the required result.

Two participants chose to use voice during the Local Deformation, although always accompanying a gestural interaction.

User Strategy

Besides identifying the gestures that were used during the tests, we are also interested in determining the type of strategy used by the users when completing our tasks. In the representation of the users' hands, we chose to increase the opacity of the hand to show the progression of the users' motion. We found an significative preference for a unimanual approach, which is directly related to the selection method that was chosen for the user tests, as described in section 3.2.1. We observed that after selecting the cylinder, participants naturally moved their hand towards the final position, in order to correct place in regard to its proper rotation and position. We also noticed that the hand used to select the object was not necessarily the dominant one, but the one that would complete the task with a single motion.

When reviewing the Twist and Bend tasks we found that most users elected to use physically plausible movements to accomplish the goal. In one instance the participant resorted to the use of voice and drawing a "C" shape. It is also worth noting that there were occasions where participants used only one a continuous motion and others resorted to quicker repetitions of that same motion.

In the Scale task we also identified a more frequent used interaction method.. To this, nineteen out of twenty participants performed two distinct actions for each axes of transformation, and the remaining participant chose to use voice to finish this task.

When examining users actions in the Local Deformation task we identified that the most frequent gesture - used by sixteen out of the twenty participants- was drawing a triangle in face of the object This action was used either alone, or with a push gesture, indicating a desire to push in the selected region. Indeed, the region selection was the most interesting observation during this task. Out of all twenty participants, only four did not perform a region selection. Of these four, two resorted pushing in with their hand directly on the object, and trying to position their hand in accordance to the shape required. The remaining two participants used a digging/carving motion to perform this task, alluding to their understanding that the object was made of a material that could be dug through.

Results and Discussions

It is clear that different tasks elicit different ways to interact with the objects. We found that a uni-manual method was preferred in the Docking and Local Deformations tasks. It has to be noticed that in the Docking task user are required to select and object and consequently, the chosen selection method had a direct influence on how the users performed that task. On the other hand, in the remaining tasks we observed a greater use of a bi-manual approach. Since these tasks represented a more accurate representation of real world hand motions, they were much more familiar to the participants and thus elicited behaviours very similar to those used for real objects. Additionally, the use of symmetrical hand motions was much more prevalent than asymmetric, although this can be attributed to the nature of the tasks itself.

It has benn possible to identify a clear set of gestures across the majority of participants. The strategies used in the experiments remained consistent for the individual tasks. This, coupled with the prevalent use of physically plausible movements indicates that users prefer to interact with the virtual objects as natural as possible.

Some users felt confused when trying to understand what were the possibilities and limitations of the system.

Concerning voice interaction, only six users used voice commands in three tasks and mostly in

two ways: (i) to determine what the system was capable of and (ii) to describe what they wanted to do or were doing with their hands. In the first case, the approach was to use specific commands to perform a specific action, like "bend right" in the case of the Bend task.

In the second case, users simply explained what they were doing.

Both the monitoring of the participants throughout the experiment and the post-experiment interviews contributed to the collection of important information regarding the user-experience. (Appendix B illustrates some details gathered from the analysis of the video recorded during the experiment)

Due to their inexperience with Virtual Reality, some users had difficulties in understanding how they could interact with the virtual objects. A user commented that the perception that the object is not real is strange at the start. However, the majority of users interacted with the objects quite easily, and appreciated the visual feedback given to their gestures. No user realized that their actions were being reproduced by the operator, and most claimed that their gestures worked like in the real world.

Most of the participants did not change the position of the objects during the tasks (excluding the Docking task) to perform them easily, instead, they chose to move themselves or tilt their heads in order to check the correctness of the operation. Overall, we retrieved important information that will be essential in continuing research into immersive shape modelling.

APPENDIX

Appendix A - Procedure to be followed for the experiment

Actual Procedure:

- Show the consent form
- Introduce the user to the experiment
    - Show images of the environment
- Teapot test with just Translation and Rotation
- Start Recording
- Perform the Tasks
- Stop Recording
- Fill questionnaire

Running the Experiment:

- Open Unity project: **Action-Specification** (currently on my github account, also on my computer and the computer in the laboratory)
- Navigate to Assets/Resources/Scenes in Unity and open **Scene 0**.
- The keyboard keys **F5-F9** are used to change the task.
- During each task, the numbers **1-3** are used to change to the different phases of the objects. Number **0** will reset the object to its current position. (All the input keys are in the InputManager.cs and TestManager.cs files).
- To select the object in any task: while the user has the hand inside the object, press **Keypad 0** to select, and **KeyPad 1**, to deselect.

Introduction to users:

- Describe all tasks (obtain the target object)
- Describe to the users what can be done:
    - Selection/Deselection of objects
        - Selection is putting hand inside the object -> Will become attached to that hand
        - Deselection is holding hand still for 2 seconds

    - Gestures with their hands
    - Voice commands

What I need to do in each test:

- **Docking Task**
    - Select / Deselect
    - Move in Editor if using Voice commands
- **Twisting Task**
    - Select / Deselect
    - Press 1-4 for the Twist
- **Bending Task**
    - Select / Deselect
    - Press 1-4 for the Twist
    - I to Invert the Rotation of the object
- **Scaling Task**
    - Select / Deselect
    - Scale in Editor
        - UI or values (UI seems easier)
- **Modelling Task**
    - Select / Deselect
    - Press 1 to change the object
    - **Do not say hole!!!**

Appendix B - Video Analysis

Summary:

- 20 participants
- Comparison of the gestures used.

Docking Task (D):

- **1 Hand: 20**
- Gesture
  - Grasping: 6
  - Open/flat : 12
  - "Italian Gesture" : 2

Twist (T):

- **2 Hands: 16**
- 1 Hand: 4
- Gesture
  - paper
  - Continuous: 10
  - Discrete: 10

Bend (B):

- **2 Hands: 13**
- 1 Hand: 6
- Voice: 2
- Gesture
  - paper
  - **Continuous: 15 (C Shape + voice)**

    - Discrete: 3

Scale (S):

- **2 Hands: 13**
- 1 Hand: 8
- Voice: 1
- Gesture
    - paper
    - **Continuous: 16**
    - Discrete: 4

Deformation:

- 2 Hands: 3
- **1 Hand: 17**
- Voice: 2 (push command only, never the full task)
- Gesture:
    - Drawing Shape Contour (Triangle)
        - Single finger : 15 (of which one hand and one finger: 14 , one finger per hand : 1)
        - Using more than 1 finger to draw and dig at the same time: 4
    - Push Operation
        - Open hand to push: 5
        - Fist to push:2
        - Hand a bit closed to push :1
        - Hand for taking away material from drawn region : 1
        - Fingers to draw and dig at the same time: 4

Video Analysis Deformation task: 20

Summary

One hand: 17

Two hands: 3

Use of voice 2 of 20 to express push command

Push command by gesture: 9 of 20 but sometimes users have been stopped saying them that the operation was completed.

Gestures for expressing deformation:

<u>Drawing shape contour</u> (triangle) :

Single finger : 15 (of which one hand and one finger: 13 , one finger per hand : 1)

Using more than 1 finger to draw and dig at the same time: 4

Push operation:

using open hand to push: 5

using fist to push:2

Using hand a bit closed to push :1

Using hand for taking away material from drawn region : 1

Using fingers to draw and dig at the same time: 4

VIDEO

2- 00:00

One hand

One finger to draw a triangle

Voice used to say at the same time “I draw a triangle”

NO clear command expressed to do depression/protusion

3 - 01:57
One hand

One finger to draw a triangle

No voice

NO clear command expressed to do depression/protusion

4- 4 - 01:10
One hand

One finger to draw a triangle

No voice

NO clear command expressed to do depression/protusion

5 - 00:10
One hand

One finger to draw a triangle

No voice

NO clear command expressed to do depression/protusion

6 - 02:40
One hand

One finger to draw a triangle

No voice

NO clear command expressed to do depression/protusion

7 - 02:00
One hand

Two fingers to draw a triangle and moving the 2 fingers like a dig

No voice

command to do depression/protusion the movement of the two fingers

8 - 00:40

TWO hands

One finger per hand to draw a triangle

No voice

NO clear command expressed to do depression/protusion

9 - 00:12

Two hands and 4/5 fingers each to draw/dig a triangle
No voice

command to do depression/protusion the movement of the the fingers

10 - 01:54
One hand

One finger to draw a triangle

No voice

Movement with ONE hand to indicate depression

11 - 01:47
One hand

One finger to draw a triangle

No voice

Movement with ONE hand to indicate depression

12 - 00:00
One hand

One finger to draw a triangle

Voice to express the willness of pushing

Two different Movements with ONE hand to indicate depression : with open hand and with fist

13 - 01:20
Two HANDS

One finger per hand to draw a triangle

No voice

NO clear command expressed to do depression/protusion

14 - 01:14

One hand

4 fingers to draw a triangle like a dig

No voice

command to do depression/protusion the movement of the 4 fingers

15 - 02:13 (file name P1160268.mov)
One hand

1 finger to draw a triangle

No voice

command to do depression the movement of one hand to extract material from the object inside the triangle

16 - 00:00
One hand

No voice

Not clear ..only one open hand to push?

17 - 01:50
ONE hand

No voice

only one open hand to push

19 - 05:57
One hand

One finger to draw a triangle

No voice

Movement with ONE open hand to indicate to push

20 - 00:35
One finger to draw a triangle

Voice to express push operation

command with one hand (not open not fist) to push on the object

21 - 00:55
One hand

First with open hand

Then One finger to draw a triangle

No voice

No command for pushing

22 - 01:25
One hand

Then One finger to draw a triangle

No voice

Movement with ONE open hand to indicate to push

Appendix B

Questionario

Appendix C – Profiling Questionnaire

Object Manipulation and Modelling

In this questionnaire we aim to determine the demographic characteristics of the users of the study.
Filling this questionnaire in its entirety should take no longer than 2 minutes.

*Campo obbligatorio

1. **ID number** *

____

2. **Gender** *

Contrassegna solo un ovale.

- ( ) Male
- ( ) Female

3. **Age** *

Contrassegna solo un ovale.

- ( ) 18-25
- ( ) 26-30
- ( ) 31-40
- ( ) 41-50
- ( ) 50+

4. **Academic Qualifications** *

Contrassegna solo un ovale.

- ( ) Basic (Until 9th grade)
- ( ) High School (10th to 12th grade)
- ( ) Bachelor Degree
- ( ) Master Degree
- ( ) Doctorate Degree

5.

How often do you use Virtual Reality systems? *

This includes Stereoscopic tabletops, Shutter Glasses, HTC VIVE, Oculus Rift, Samsung Gear VR, etc.

Contrassegna solo un ovale.

- Never
- Less than one time per month
- At least once per month
- At least once per week
- Daily

6.

How often do you use Gesture Detection system? *

This included the Microsoft Kinect, HTC VIVE Cameras, Leap Motion, etc

Contrassegna solo un ovale.

- Never
- Less than one time per month
- At least once per month
- At least once per week
- Daily

7.

How often do you use 3D Modelling Programs? *

Such as Blender, 3DS Max, Maya, etc

Contrassegna solo un ovale.

- Never
- Less than one time per month
- At least once per month
- At least once per week
- Daily

8.

How often do you play video games? *

Contrassegna solo un ovale.

- Never
- Less than one time per month
- At least once per month
- At least once per week
- Daily

Recent titles from the IMATI-REPORT Series:

2017

17-01: *BPX preconditioners for isogeometric analysis using analysis-suitable T-splines,* D. Cho, R. Vázquez.

17-02: *Initial-boundary value problems for nearly incompressible vector fields, and applications to the Keyfitz and Kranzer system*, A. P. Choudhury, G. Crippa, L.V. Spinolo.

17-03: *Quantitative estimates on localized finite differences for the fractional Poisson problem, and applications to regularity and spectral stability*, G. Akagi, G. Schimperna, A. Segatti, L.V. Spinolo.

17-04: *Optimality of integrability estimates for advection-diffusion equations*, S. Bianchini, M. Colombo, G. Crippa, L.V. Spinolo.

17-05: *A mathematical model for piracy control through police response*, G.M. Coclite, M. Garavelllo, L.V. Spinolo.

17-06: *Uncertainty Quantification of geochemical and mechanical compaction in layered sedimentary basins*, I. Colombo, F. Nobile, G. Porta, A. Scotti, L. Tamellini.

17-07: *VVS medical workflows: definition of a static workflow for part-based annotation of wrist bones & web service oriented architecture for executable workflows*, M. Pitikakis, F. Giannini.

17-08: *Computational Methods for the Morphological Analysis and Annotation of Segmented 3D Medical Data*, G. Patanè, F. Giannini, M. Attene.

17-09: *Applying Functional Principal Components to Structural Topology Optimization,* G. Alaimo, F. Auricchio, I. Bianchini, E. Lanzarone.

17-10: *Convergence of Sparse Collocation for Functions of Countably Many Gaussian Random Variables (with Application to Elliptic PDEs),* O. G. Ernst, B. Sprungk, L. Tamellini.

17-11: *Soluzioni Open Source per la scalabilità di servizi di raccolta ed utilizzo di dati ambientali da sensore*, A. Quarati, A. Clematis, V. Levati.

17-12: *Analysis of a multi-ring technique for 3D pattern recognition*, E. Moscoso Thompson, S. Biasotti.

17-13: *A sparse-grid isogeometric solver*, J. Beck, G. Sangalli, L. Tamellini .

17-14: *On the singular local limit for conservation laws with nonlocal fluxes,* M. Colombo, G. Crippa, L.V. Spinolo.

2018

18-01: *Arbitrary-order time-accurate semi-Lagrangian spectral approximations of the Vlasov-Poisson system*, L. Fatone, D. Funaro, G. Manzini.

18-02: *A study on 3D shape interaction in virtual reality*, E. Cordeiro, F. Giannini, M. Monti, A. Ferreira