Grasping 3D Space

Using Physical Shapes as Representatives for Virtual Objects

Alexander Wirtz, Marcella Klauser

Interaction Engineering

Winter semester 23/24

Supervisor: Prof. Dr. Michael Kipp

Augsburg University of Applied Sciences

Hochschule Augsburg
University of Applied Sciences

Table of Contents

1. Abstract

When looking up online tutorials for various 3D software, chances are the first part of many of the various series available will be about camera control and navigation in 3D. Given how humans navigate 3D at basically all times in the real world, it might seem a surprise how this needs to be specially addressed. The problem lies that in the real world us humans have access to a great number of joints and limbs to control, whereas when working on a computer, movement is generally limited to the two-dimensional plane of a mouse pad. These two degrees of freedom do of course not cover the six required to have complete control over 3D movement, consisting of movement in the three dimensions, as well as rotation around each of these axes. To combat this, most programs offer control over two axes at a time, such as panning along the view plane and at the press of a button, rotation around a virtual point. Which button to press however, is not conventionalized and often varies from program to program, thus creating a need for the aforementioned video tutorials on camera control.

The idea behind this project was to create an alternative input method which allows simultaneous control over all six degrees of freedom. This is achieved by tracking a tangible object using ArUco markers, the rotation and position of which are interpreted using Python and finally applied to a Blender viewport to control the camera. This allows for intuitive usage as a virtual object on the screen will match the movements of the tangible, which can be freely handled by the user. This very direct approach to mapping is supposed to especially help inexperienced users.

2. Motivation

As stated in the abstract, every 3D software uses a slightly different kind of key and mouse kombinations to perform different manipulations in 3D space. This makes it particularly difficult for newcomers to learn a program quickly. This problem also arises when people decide to switch from one software to another. They find it difficult to adapt to the new keyboard combinations as they are used to different ones. There exist computer mouses that are especially designed for 3D applications, however they are very expensive and complex to use.

To solve all these problems we invented a method to use a tangible for 3D navigation in the most easy and intuitive way possible. With a tangible, the user can perform all six degrees of freedom as he is used to in the real world. A webcam tracks the rotation and movement of the tangible. The viewport of the 3D software blender will then mirror the movement. For the shape of the tangible we decided to use a cube.

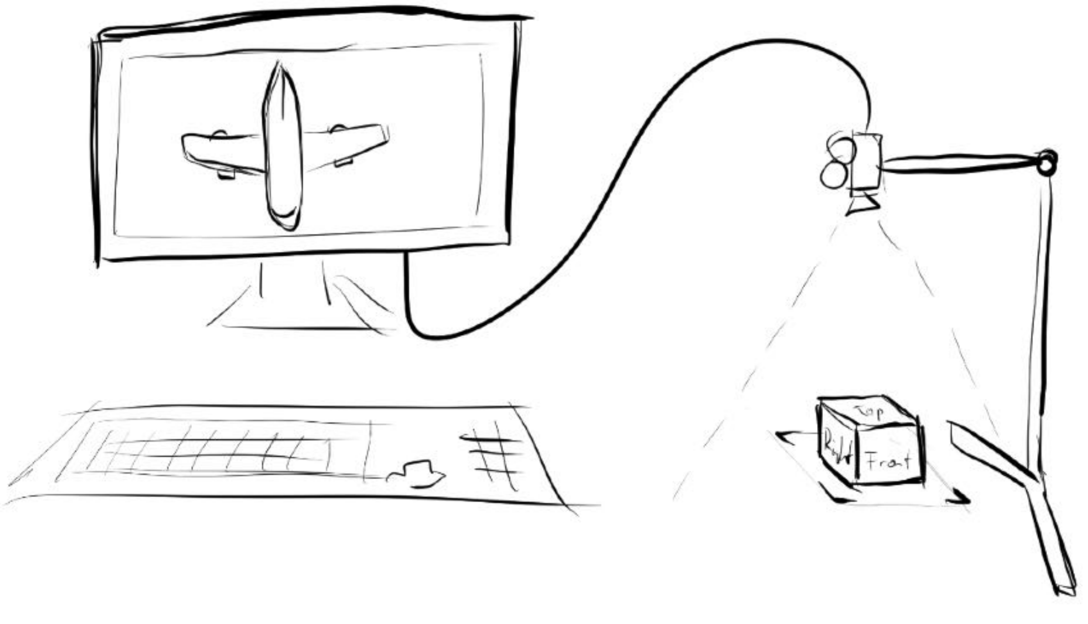


Image 1: Sketch of the prototype

3D navigation with a tangible can become handy for presentations. This allows the speaker to keep facing the audience while manipulating the object he wants to show. Another use case is to have an easy and intuitive way for beginners to learn 3D navigation. Furthermore it can enhance the workflow when working in 3D. For example while setting up a camera in blender to render an animation, it can become much easier to do this with the tangible, than with the mouse. With the tangible, the user can rotate and move the camera in one flow, while with the mouse he has to do the rotation and the translation separately.

There already exists related work on this topic. Sanmathi Dangeti, Chunhui Zheng and Yingjie (Victor) Chen already examined an object-in-hand tangible user interaction for navigation of 3D objects in modeling. They use an iPhone as the tangible object and try to detect the movement with the built-in accelerometer.
The problem with an accelerometer is that it can detect horizontal rotation well, but can’t really measure around a vertical axis. That's why we decided to use a cube as the tangible with ArUco markers for the movement detection.

Image 2: Setup of the existing paper
Source: https://dl.acm.org/doi/abs/10.1145/2839462.2856555

3. Concept

“Grasping 3D Space” is a tangible object which is used to intuitively navigate in 3D space. The user can rotate and move the object around in space and the viewport of the 3D program blender will mirror that movement. On the tangible are ArUco markers. A webcam tracks the position and rotation of the markers and translates them via OpenCV and a Python script into blender. The goal of this project is to create an easy to use technique for 3D navigation.

There are basically two different types of interaction. First, the user is able to roll, yaw and pitch the tangible in every direction. The viewport of blender then exactly mirrors the movement.

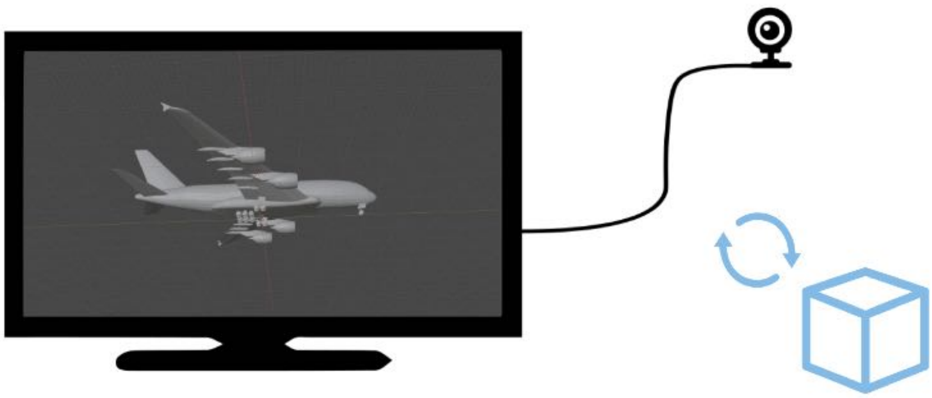


Image 3: Rotation of the tangible

As for the second interaction it is also possible to move the tangible along the x-, y-, and z-axis in space. The movement is also translated directly to the viewport.

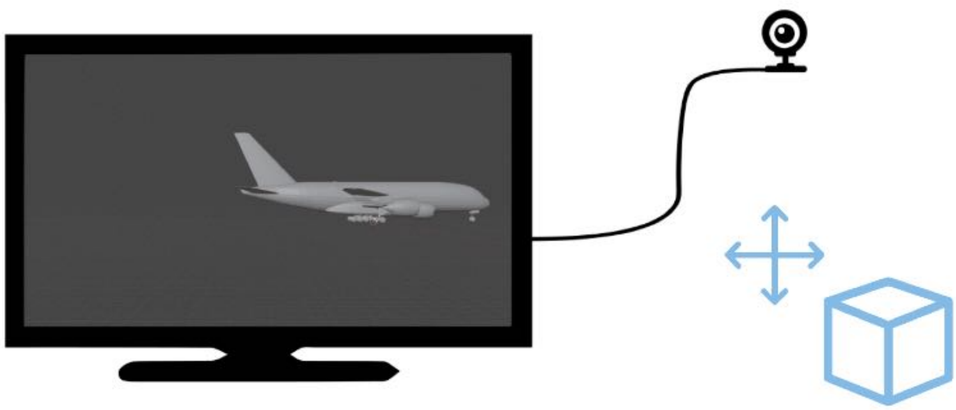


Image 4: Translation of the tangible

4. Implementation

The tangible was created by 3D-printing a cube of 6 cm side length and attaching 5 cm side length ArUco markers printed on basic paper using double-sided duct-tape. The technical setup consisted of only a laptop with a built-in webcam. Using OpenCV the webcam detects the markers and estimates their position and orientation. To account for the different coordinate system used in 2D where the origin of the image starts in the top left corner as opposed to the center, the location of all markers is shifted across the length and width of the image.

The location can simply be averaged between all detected markers and, while technically not entirely mathematically correct, passed into Blender with relative ease. During initial testing the gain factor for the translation was reduced quite substantially as the camera would otherwise often go flying across the screen at the slightest movement. Handling the rotation was a lot more complicated. Based on the ID of the marker, its orientation is transformed to make all six possible sides of the cube uniform. This way the algorithm can determine the absolute rotation of the tangible regardless of which side/combination of sides is visible. This requires all tangibles to use the exact same layout of markers.

To accommodate for the user facing camera, some of the resulting transposition and rotation values had to be inverted. Otherwise problems such as the viewport rotating in the opposite direction of the tangible, creating a sort of cog movement. Simple trial-and-error sufficed to find the correct matrix entries to modify. The final matrix is then passed into a smoothing filter to reduce jitter and then set as the view matrix of the viewport.

5. Evaluation

The goal of the user test is to validate the hypotheses: “A tangible simplifies the navigation in three-dimensional space”. Therefore the tangible is compared with a common computer mouse.

Seven participants were tested and interviewed for the evaluation. The age of the test subjects ranged from 22 - 29 years. Two of them had no previous experience with navigation in three-dimensional space, while the rest are slightly to well experienced with 3D editing softwares or other applications which rely on 3D navigation.

5.1 Test procedure

Participants were given increasingly difficult tasks to complete with both the tangible and a regular mouse. The order of input devices was randomly chosen for each participant. Before the first task the controls were explained to each participant and they had 2 minutes to do a warm-up practice for each device. Afterwards the participant had to complete the following three tasks. For each task, the test conductor would record the task completion time. Afterwards they had to fill out a questionnaire. The full questionnaire can be seen in the appendix.

For the first task participants were given a virtual scene with a cube which had the numbers zero through 5 written on each side. The conductor would read out a string of ten numbers and the participant had to rotate the scene to show the desired number frontally in the center. The numbers read out were the same across all participants.

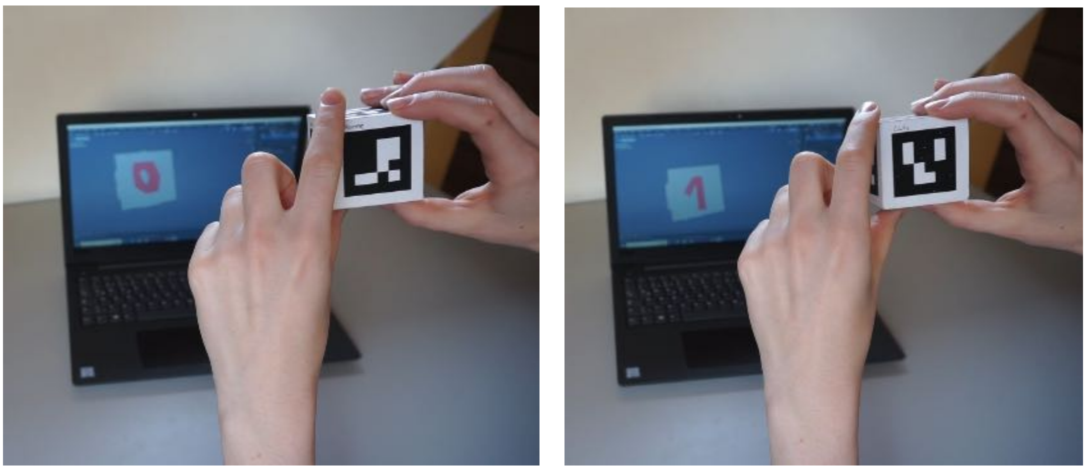


Image 5+6: Task 1

Task two added translation onto the existing challenge of rotation. This task featured a 3D model of a large civil airplane. The participants were dealt a set of five flipped cards, each one had a depiction of said airplane in a distinct location on it. The participants would then unflip any of the five cards and had to recreate the angle and position shown on the card. There was unfortunately no way to objectively judge when the participant had completed the task, so the test conductor had to manually decide when the task was completed successfully.

Image 7: Different viewport settings for Task 2

Image 8: Task 2

The final task tried to emulate a workflow scenario where the virtual scene contained a plain cube and an abstract, more complex shape next to it. The task was to recreate the complex shape, which was possible by only using the extrude command on all 6 sides of the cube. For this, the mouse and keyboard had to be operated simultaneously alongside the tangible.

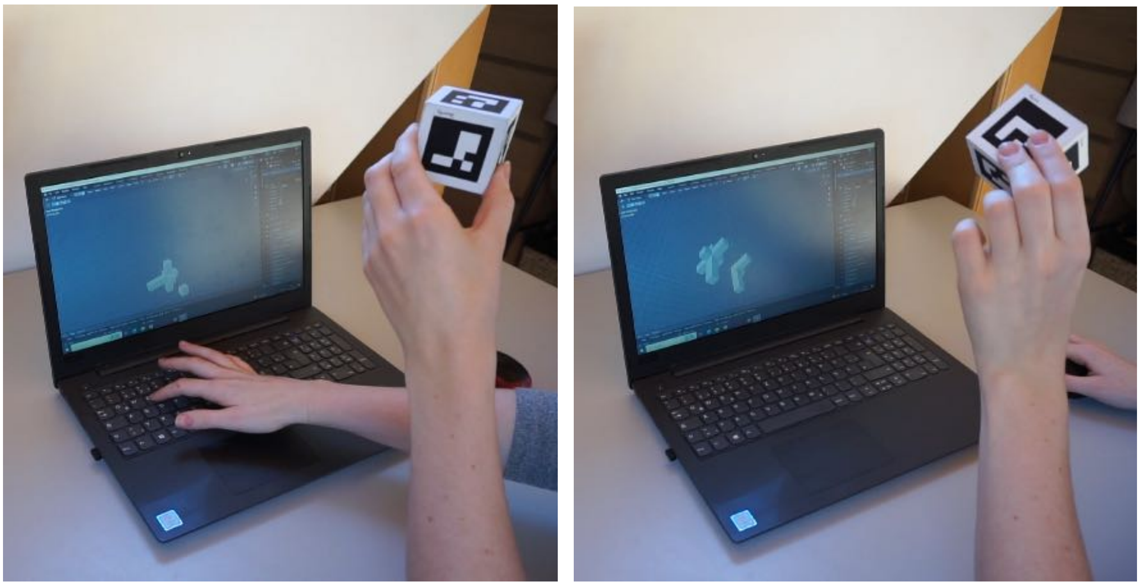


Image 9+10: Task 3

5.2 Results

The tangible performed rather poorly in comparison to the mouse. On average, tasks were completed 54%* slower with the tangible. Especially for the third task, in which the tangible should be integrated into a typical workflow of blender, the participants took almost twice as long with the tangible than with the mouse. This can be seen in the figure below. However inexperienced users took 29% longer than experienced users when using the mouse, but only 15% longer when using the tangible. In addition, five out of seven users found the tangible to be fatiguing and three out of seven found it frustrating to use.

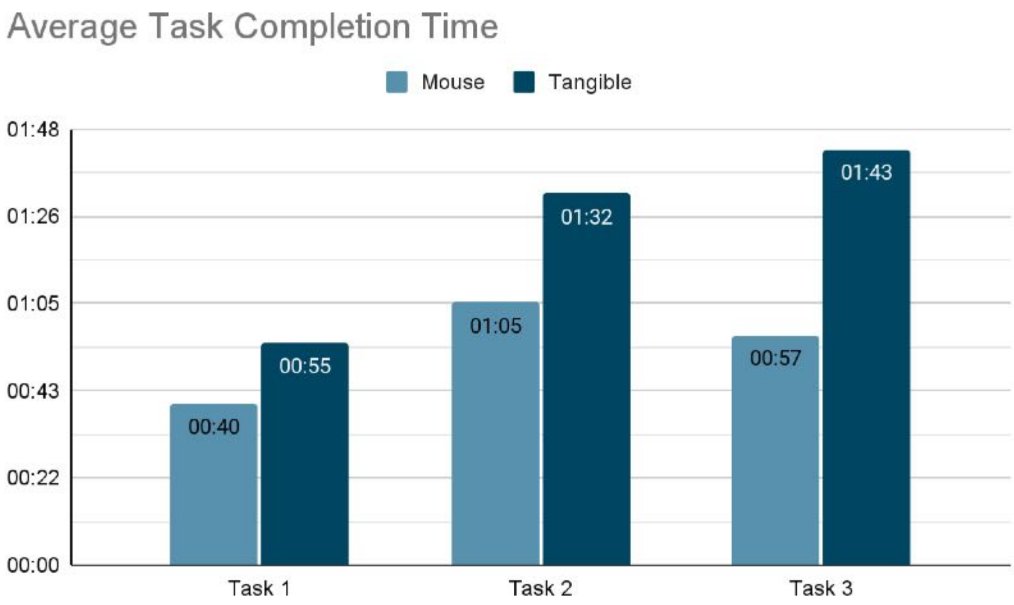


Image 11: Average Task Completion Time

This is the result of various small issues adding up, the probably most important one being the lack of a dedicated clutch function, which had to be left out due to time and software restrictions. While it is technically possible to engage a clutch and lock the viewport in place by quickly removing the tangible from the detection area of the camera, this unintended behavior was not told to or used by the participants. In combination with the usage of a built-in laptop camera, which, due to the angling of the screen, was partially pointing upwards, required the testers to mostly keep their arm in the air and left them unable to rest their elbow on the table. Another, in hindsight obvious, issue was the occlusion problem. The 5 cm markers on the 6 cm cube left only half a centimeter from the edge of the cube to the edge of the marker. This made them very easy to accidentally occlude, as the marker detection would break with only a tiny part of the marker not visible to the camera. For further iterations a larger cube and/or smaller markers are recommended. To end on a positive takeaway, the tangible was judged comparably intuitive to a mouse and deemed more fun to use by the participants.

6. Conclusion

6.1 Summary

Our hypothesis “A tangible simplifies the navigation in three-dimensional space” could not be confirmed by the user test. Not only did the participants take more time to complete a task while using the tangible, the majority of them still preferred the mouse over the tangible. Nevertheless, the user found it to be a fun and intuitive way to navigate in 3D space. We are convinced that with an improved prototype and a longer familiarization period, better results can be achieved. To confirm this assumption, further user tests must be conducted.

6.2 Future Work

As shown in the user test, a clutch should be added to lock the position of the tangible. It would also be possible to attach the tangible to a telescopic arm, which would automatically provide a clutch. With that solution, the aspect of fatigueness would become obsolete, as you would no longer have to hold the tangible in your hand all the time. Furthermore, attempts should be made to use an external webcam. In the user test, the built-in webcam of the laptop was used. As a result, the screen was difficult to see as the tangible always had to be held directly in front of it, making the workflow more difficult. For further tests, it can also be investigated whether a different shape or type of tangible (e.g. a sphere) can provide better results.

Appendix

Questionnaire

Name *

Meine Antwort

Age *

Meine Antwort

Gender *

- ○ male
- ◉ female
- ○ other

Do you have experience with 3D editing software or other applications which rely on 3D navigation (e.g. games)? *

	1	2	3	4	5	
none	○	○	○	◉	○	frequent user

If yes, with which software?

Meine Antwort

I think it is easy to navigate in 3D *

	1	2	3	4	5	
strongly disagree	○	○	○	○	○	strongly agree

I think the usage of the system is fatiguing *

	1	2	3	4	5	
strongly disagree	○	○	○	○	○	strongly agree

I have to practice before I can use the system correctly *

	1	2	3	4	5	
strongly disagree	○	○	○	○	○	strongly agree

I think this system was intuitive to use *

	1	2	3	4	5	
strongly disagree	○	○	○	○	○	strongly agree

I think that I would like to use this system frequently *

	1	2	3	4	5	
strongly disagree	○	○	○	○	○	strongly agree

I felt frustrated using this system *

	1	2	3	4	5	
strongly disagree	○	○	○	○	○	strongly agree

I think it is easy to navigate in 3D *

	1	2	3	4	5	
strongly disagree	○	○	○	○	○	strongly agree

I think the usage of the system is fatiguing *

	1	2	3	4	5	
strongly disagree	◯	◯	◯	◯	◯	strongly agree

I have to practice before I can use the system correctly *

	1	2	3	4	5	
strongly disagree	◯	◯	◯	◯	◯	strongly agree

I think this system was intuitive to use *

	1	2	3	4	5	
strongly disagree	◯	◯	◯	◯	◯	strongly agree

I think that I would like to use this system frequently *

	1	2	3	4	5	
strongly disagree	◯	◯	◯	◯	◯	strongly agree

I felt frustrated using this system *

	1	2	3	4	5	
strongly disagree	◯	◯	◯	◯	◯	strongly agree

Open Questions

Which system did you prefer? *

- ◯ Computer Mouse
- ◯ Tangible

Why do you like this system more?

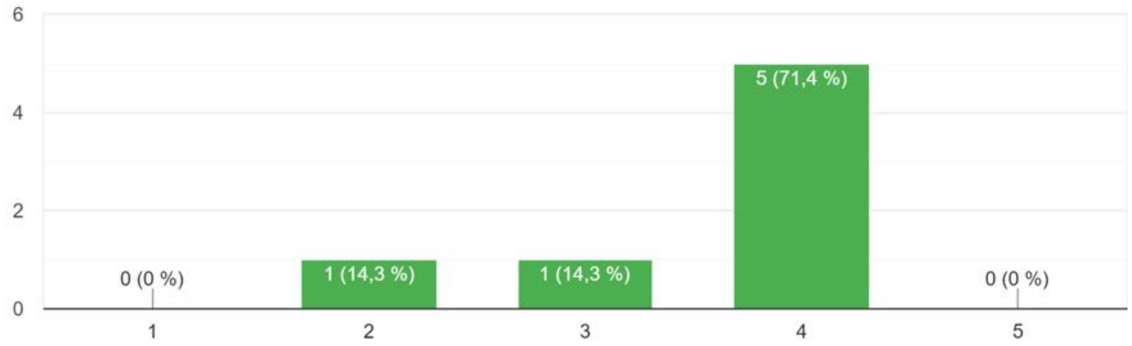
Meine Antwort

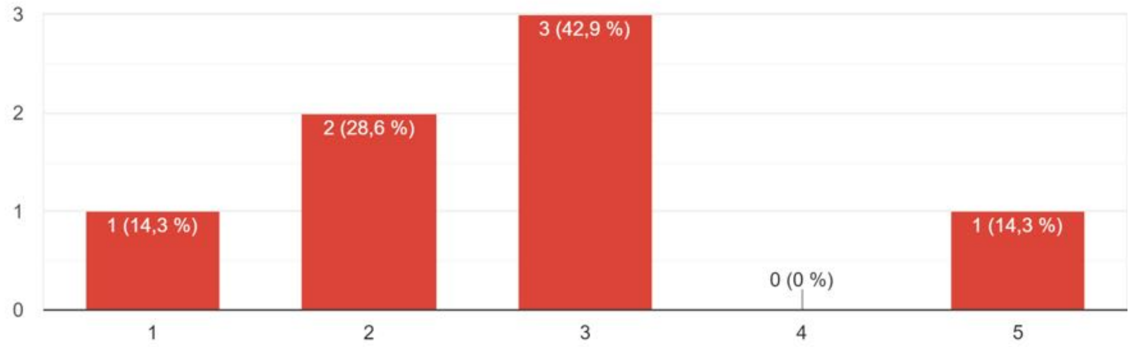
What do you think could be improved when navigating in 3D space?

Meine Antwort

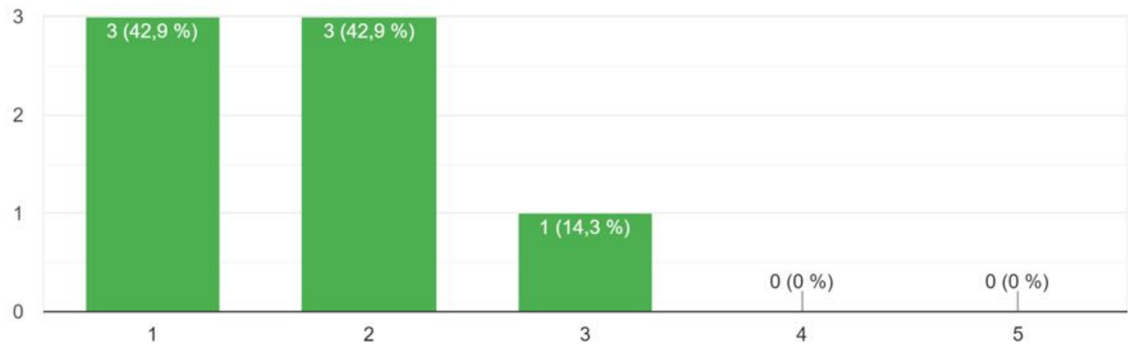
Answers of the questionnaire comparing mouse (green) with the tangible (red)

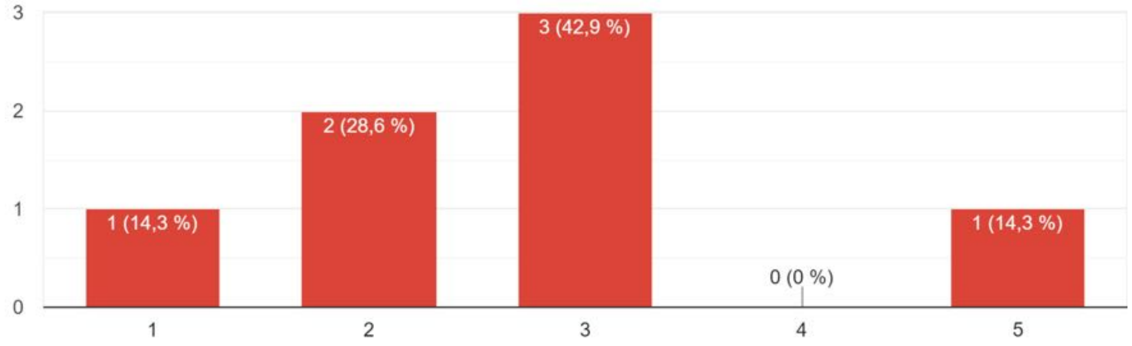
1. **I think it is easy to navigate in 3D**



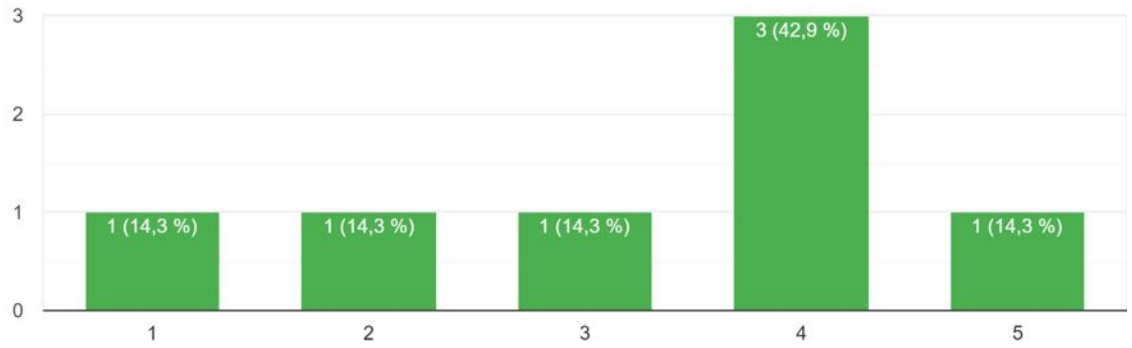


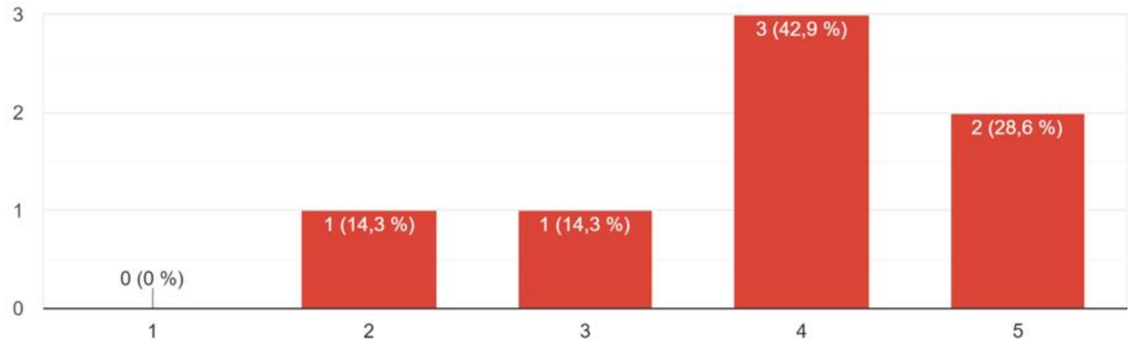
2. **I think the usage of the system is fatiguing**



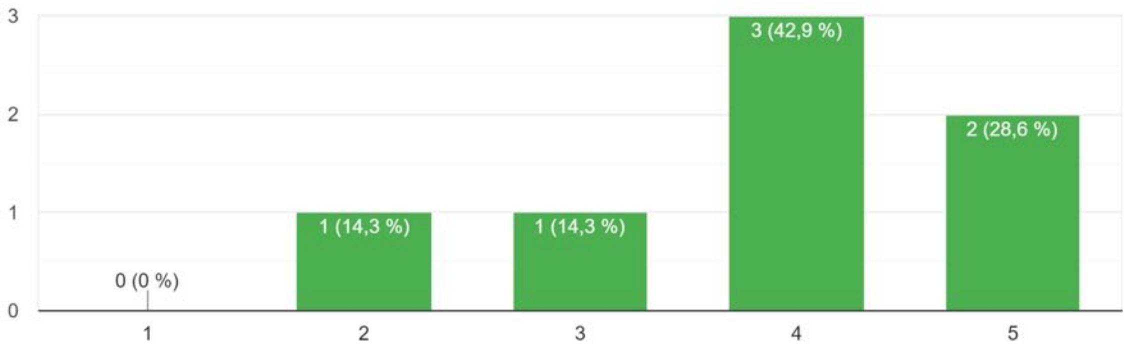


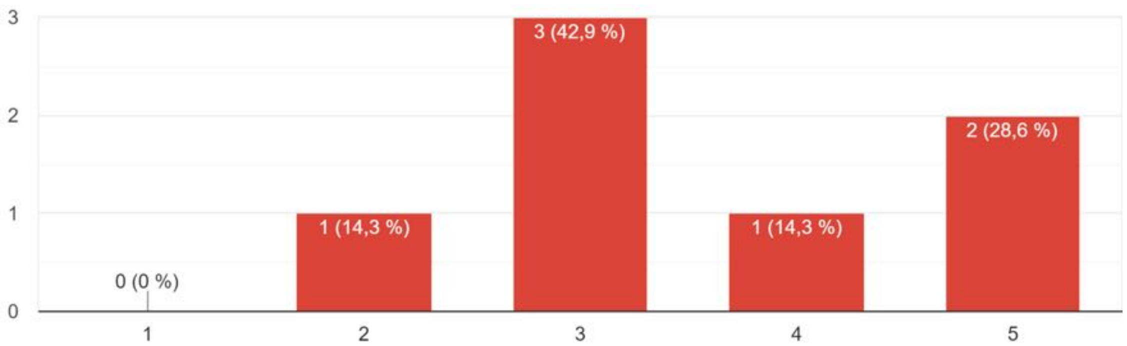
3. **I have to practice before I can use the system correctly**



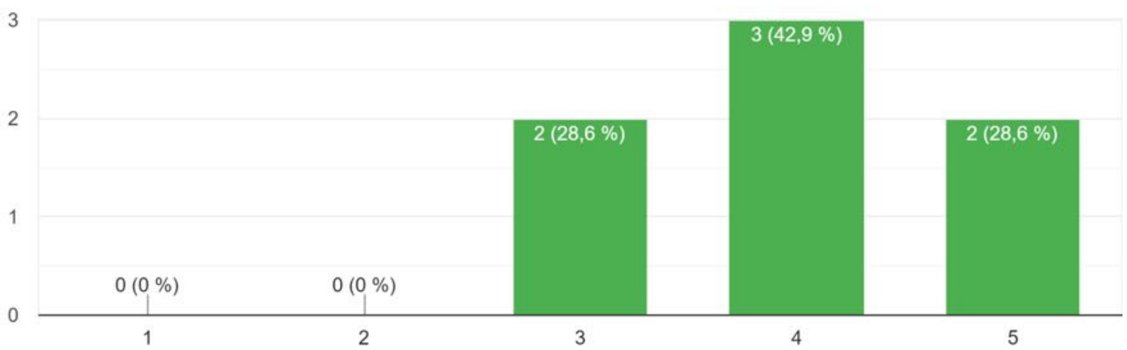


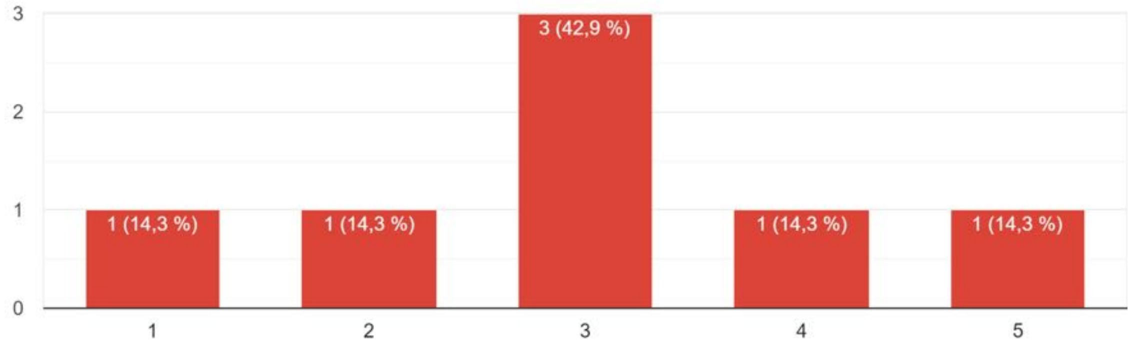
4. I think this system was intuitive to use





5. I think that I would like to use this system frequently





6. I felt frustrated using this system