Evaluating the Effects of Fidelity in Collaborative Object Manipulation

Sachin Bharambe[*] Virginia Tech.
Amruta Kulkarni[†] Virginia Tech.
Surabhi Gaopande[‡] Virginia Tech.
Varun Subramanian[§] Virginia Tech.
Pranav Murthy[¶] Virginia Tech.
Nicholas Polys[‖] Virginia Tech.

Abstract

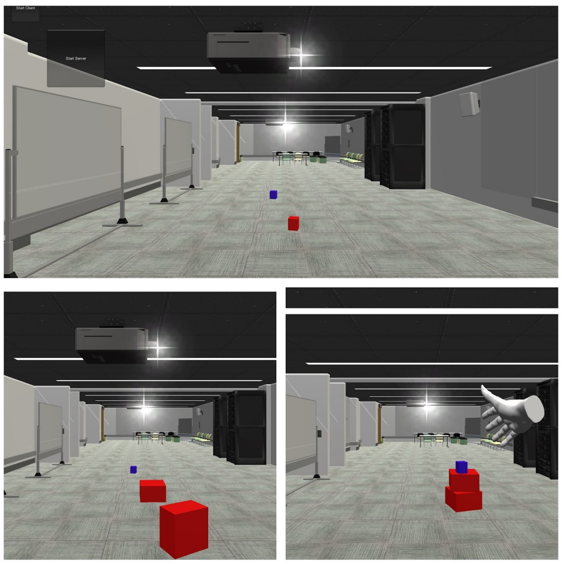

Figure 1: Collaborative Virtual environment

Globalization and the advancements in the technological world, pose the need for designers working across different geographic regions to work collaboratively on 3D content, especially in fields such as urban planning, engineering or architectural design. This paper presents our work on the design and evaluation of a collaborative 3D user interaction platform for object arrangement, where remote users manipulate virtual objects by gestural hand-based interactions. The platform was created to provide the remote users a realistic experience, offering higher levels of fidelity which facilitate seamless remote collaboration. One pressing question in the field of Virtual Reality is how much fidelity is enough? The level of visual display fidelity has many facets ranging from field of view (FOV), field of regard (FOR) to the display size and resolution to the head-based rendering (given by accurate head tracking). We tested a high-fidelity manual interface of the Leap Motion Controller with the Oculus Rift with enhanced head tracking capability (higher Field-of-Regard), a desktop display (limited Field-Of-Regard), varying the presence of shadows and the presence of rigid body physics. The goal of the object manipulation task was to replicate a target assembly of virtual cubes (figure 2). Two users had to collaborate on either side of a virtual workbench to achieve the target block configuration. We evaluated the system against the metrics of time, accuracy and user satisfaction. Our work demonstrates the challenges and opportunities in using high-fidelity cues for providing seamless interaction for remote collaboration.

Keywords: collaborative interaction, gesture input, 3D Interaction, user studies.

1 Introduction

Remote collaboration is a challenging research topic, with many researchers innovating ways to tackle a thorny problem. Remote collaboration is more and more essential as a global workforce may have designers located in distant geographical locations. While simple audio or video conferencing technology such as Skype or Google Plus Hangout might be sufficient for communicating with certain people and on certain topics, these tools do not provide professionals with a shared task space where they can collaborate on more complex tasks, such as an architectural design meeting in which these experts must be able to visualize digital content at various level of details and modify the content. Many of these remote collaboration tasks are complex and may benefit from high-fidelity 3D User Interfaces. Virtual Reality (VR) and high-fidelity 3D User Interfaces are a promising area for such applications which require user interaction for collaborative visualization tasks. Not only do 3DUIs allow natural interaction between the real and virtual environment, they also enable the delivery of spatial and gestural cues of remote collaborators. The recent advancements in VR have taken significant strides with respect to the growth of the display technologies, especially head mounted displays (HMDs) . Many exciting display and sensor technology are today available both to the academic community and the general public, allowing researchers to focus on the interaction aspect of designing a VR-based user interface instead of reinventing the tracking capability or display technology.

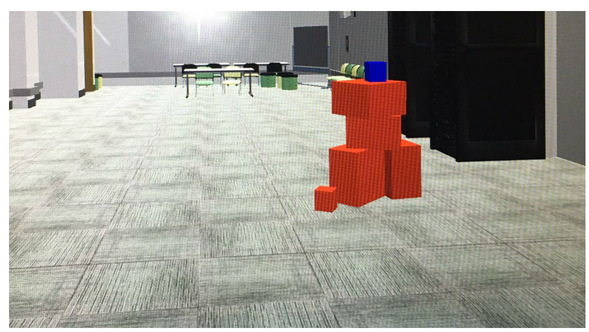

Figure 2: Target assembly

In this work, we develop a collaborative platform for users to prototype their designs using 3D interactive hardware and software

*e-mail: sachin77@vt.edu
†e-mail:amrukk6@vt.edu
‡e-mail: surabhig@vt.edu
§e-mail:varuns92@vt.edu
¶e-mail:pranavvm@vt.edu
‖e-mail:pranavvm@vt.edu

tools. Being a VR based system, our platform aims to provide a seamless interface between the virtual and the physical by placing virtual objects in the real world that the user can perceive in space by wearing a binocular optical head-mounted display like the Oculus Rift. The interactive 3D constructive solid geometry and object assembly task will be achieved through the use of Leap motion sensor for gesture recognition. In addition, users can simultaneously view and interact with the virtual objects naturally by using their hands. Our system also supports remote collaboration, where remote users hands are shown to increase the gesture and action awareness among collaborators.

The goal of our object manipulation task was to build the tallest possible tower with the virtual cubes in given time limit. For our trial studies, we assigned the task of assembling a tower of maximum height within a fixed time limit and evaluated participants based on the metrics of time, accuracy and satisfaction. Two (can be extended to more) people present in the a virtual workbench would collaborate to achieve this common goal. We vary independent variables such as field of regard (Oculus Rift vs Screen display), complexity of target assembly (using shadow effects), and presence of physics in the virtual environment and evaluate the effectiveness of the system based on the metrics of time, accuracy and satisfaction to replicate the target assembly.

This paper is structured as follows: Section 2 discusses the related work with respect to our proposed system. Section 3 discusses the design principles on top of which our system operates and is built. The system architecture and the software platform are discussed in detail. This chapter also discusses both the hardware context and the software context of the system. Section 4 discusses the implementation details of our system and the experimental setup for our evaluation. Details related to System setup and calibration procedures are also outlined in this chapter. Section 5 discusses the user study and the results that we obtained. Finally, Section 6 provides a summary of the work and outlines future work.

2 Related Work

Extensive research in this field has been carried out with the intention of evaluating 3D input devices in terms of 3D interaction techniques and its relation to user performance. As a result of constant development of interaction devices and rendering systems, research in this field is still a common practice.

In a related work, a study by Joann et al on Pointing Task Evaluation of Leap Motion Controller in 3D Virtual Environment [3] provides a good evaluation of 3D pointing tasks using Leap Motion sensor to support 3D object manipulation. In this paper three controlled experiments were performed in the study, exposing test subjects to pointing task evaluations and object deformation, measuring the time taken to perform mesh extrusion and object translation. Qualitative data was gathered using the System Usability Scale questionnaire. Their data reveals a strong correlation between input device and performance time suggesting a dominance of the Leap Motion gestural interface over mouse interactions. Based on their results from the experiments, they concluded that presented 3D input device, Leap motion outperformed mouse interaction only in single target situations, showing that 3D translation is less cumbersome when the z axis is provided as input based on real-life movement mappings.

Our experiment was paralleling the work accomplished by Joann et al [3], as our research also evaluated the performance of Leap motion for 3D object modeling (not object manipulation) in a virtual environment. Similar to the researchers in this paper, we exposed our test users to a choice of controlled experiments that involved object modeling with and without gravity, with and without Oculus Rift and with and without shadows and the evaluated the data collected during these controlled experiments. Our experiments revealed that our test subjects preferred Leap motion controller as an input device over Kinect for virtual reality experiments due to the intuitive feeling it induces in a user when using it for object manipulation in a virtual environment.

A relevant work by Otmar Hilliges et al, on HoloDesk: Direct 3D Interactions with a Situated See-Through Display [9]. The authors presented a novel device called as HoloDesk which has the ability to capture user gestures which would in turn be translated to actual object manipulation that could be viewed in a see-through screen incorporated as a part of the device. The system employs Kinect sensors to record user gestures. The paper then describes various techniques and gestures involved in that process, Similar to the experiment carried out by the authors of this paper, we performed our experiment targeting a 3D object modeling and interaction in a 3D virtual environment.

The task for the users was to stack 3D block on top of each other to obtain the highest stack of blocks possible, before these blocks tip over. In our experimentation, the users gestures were enabled using an avatar in the form of an arm that replicated the users arm gestures and wrist actions. We evaluated the performance of our users with the help of a Leap motion controller instead of a Kinect and evaluated the performance of these users with and without an Oculus Rift. Our experiments also examine the relation between users gender, video game experience, 3D software and 3D input device experience and their performance with the 3D User Interface.

3 Exposition

Our collaborative system is made up of multiple components: the Server (manages communication for all clients), Unity 3D clients running, an input device for 3d body tracking, and a 3D display. The Servers primary purpose is to distribute the information received from one client to all the other clients. While working on this design, we experimented with different inputs devices like Kinect V2 and the Leap Motion. But for our experiment, which is more focused on hand gestures, we selected the Leap Motion as the choice input device for our system. The Leap Motion is used to extract pose of the hand and recognize hand gestures. The visualization of the model and environment is done by either the Oculus Rift (high FOR) or a laptop (low FOR).

3.1 Interactive Set-Up Methodology

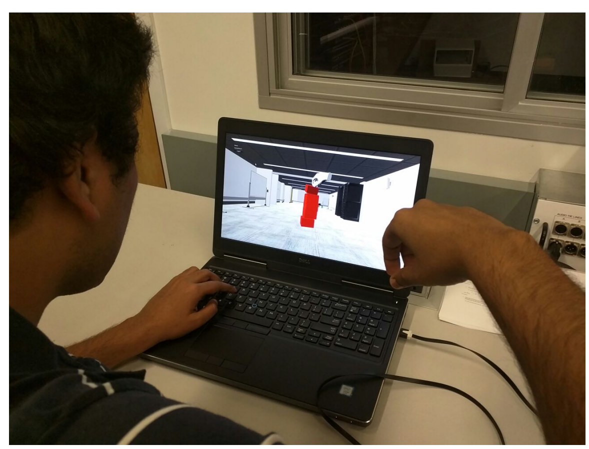

Figure 3: User study without Oculus

Figures 3 and 4 show the Oculus and the laptop conditions we tested. The task given to these users was to construct a tallest tower with collaboration of remote user in virtual work place. User can navigate to any corner of the room to collect and arrange the cubes using four keyboard keys. To provide and high-fidelity interaction for performing the task of object manipulation, our system recognizes pinch and drop actions of the hand.

Figure 4: User study with Oculus

3.2 Stimuli

3.2.1 Platform

We used the Unity game engine as the software platform since it offers a wide range of benefits including cross-platform support, abundant technical support, support for Physics, and availability of whole range of basic models which can be used to construct relatively complex virtual worlds. The setting of the tasks was a 3D model of our Moss Arts Center building on campus. The virtual cubes and other virtual objects that we used in this project were created and rendered using Unity. The project involved:

- Creation of the virtual objects and corresponding virtual ambiance,
- Adding the physics of interaction with the other objects like collision detection, rigid body etc.,
- Adding the physics of gravity, which determines the nature of interaction and behavior of the object in the virtual word.
- Controlling the object with respect to Leap motions input.

3.2.2 Network and Connection

Real-time collaboration between two users is a key component in this project. This involved establishing a connection between two users (systems) over the network using WebSocket, which uses an underlying TCP connection to communicate between the server and the clients. In this project, we have created a central server component using Tornado server in Python. The Tornado server is a Python based Web framework and asynchronous library for networking that employs non-blocking network I/O and provides the framework essential for establishing a WebSocket. This server is responsible for communicating the state of objects between the collaborating user systems. The server acts as a bridge between the two users, by acquiring the state (position) of the cubes from each user environment and then communicating it to the other user. This helps in synchronizing the position of the object positions in real time across the two users, achieving a seamless collaboration effect with changes to model getting reflected in real time.

3.2.3 Hardware Design

In this remote interaction system, one user has a system which will act as server and other users have systems running the clients. The cube objects are then spawned in client and server sides. The Leap Motion client captures the hand motion over it and sends the data to Unity interface which accepts the data and then translates the hand motion into virtual hand motion and in turn moves the cube. The position of the cubes is then is transferred to the server, which in turn forwards the coordinates of to the other instance of the program over the network. This act synchronizes the position of the cubes between the instances, leading to real time update and hence the collaboration. The users can also add Oculus to their computers, which will enable them to view the virtual world in stereo, thus offering a 3D perspective which could facilitate a better interaction with the objects in the virtual world.

4 Evaluation

We conducted a user study to test the usability and effectiveness of the system as a gestural remote collaboration system on virtual content. Subjects (13 males, 7 females) were invited to participate in a formal user test in a controlled environment in which they have to perform specific tasks in pairs and provide feedback to structured questions.
Each user was given 20 minutes to complete the task. Users were given a reference assembly model and asked to mimic its assembly in the virtual collaborative environment. The users were asked to use different tools for each test case and the time required to complete the task in each case was used as a performance metric. Different target models were built, varying the presence of physics and shadows. In this test case we analyzed the response time in different work models. By employing concepts of physics, and imparting shadows to the macros in the assembling objects the difficulty level and thrill is further enhanced. Post the system interaction, the participants were then asked to fill out a post experiment survey in which they had to give ratings to the overall system performance based on different metrics. The detailed questionnaire was then used to analyze the overall system performance.

5 Results and Analysis

5.1 Quantitative results

The correlation between various parameters such as users gaming experience, 3D software experience, experience using 3D interaction devices, Purdue Spatial Visualization Test [4] scores with their over- all performance score in the VE is calculated using the Pearson Product-Moment Correlation Coefficient. If the value of R is close to +1, this indicates a strong positive correlation, and if R is close to -1, this indicates a strong negative correlation. From the figures, it can be inferred that the there is quite a strong correlation (R = 0.587) between the users' Purdue Spatial Visualization test score and the rating given by the users for the overall experience with our system(Figure 5). The correlation between the users' experience with 3D software or video games and the overall rating given by them is not very strong, but is still positive hence significant(Figure 6 and Figure 7).

Users were asked to rate their ease with using the Leap Motion controller on a scale of 1 to 7. The mean is 4.2 and standard deviation is 1.66. (Figure 8). User ratings for the preference of Oculus for the task of assembling objects have a mean of 5.19 and standard deviation of 0.95 (Figure 9). User ratings for the preference of physics for completion of the task have a mean of 5.8 and standard deviation of 1.12 (Figure 10). User ratings for the value of shadows for completion of the task have a mean of 4.95 and standard deviation of 0.97 (Figure 11).

While the users were completing the task of constructing towers with the blocks, the number of cubes assembled and the time

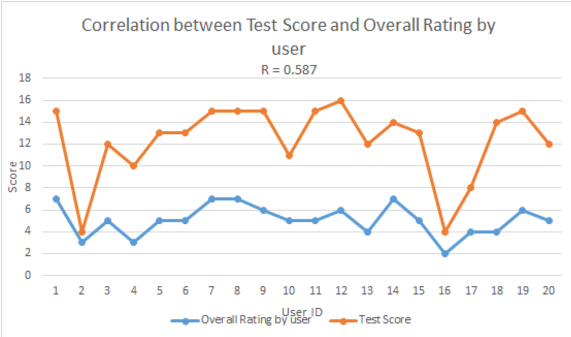

Figure 5: Correlation between Purdue Spatial Visualization Test[4] Score and Overall rating by user R = 0.587

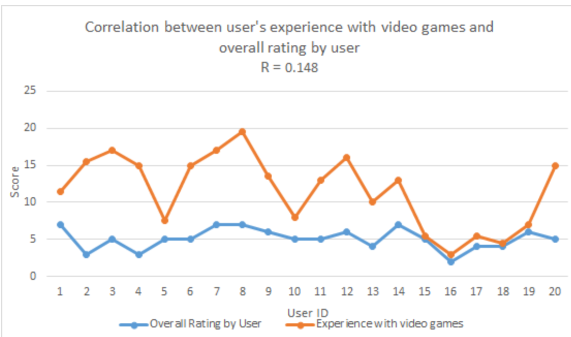

Figure 6: Correlation between user's experience with video games and overall rating by user R = 0.148

taken for each user in each scenario was recorded. A two-by-two ANOVA was performed for cube manipulations per minute with/without Oculus and with/without Physics. The results indicate that the effect of Oculus is statistically significant with a p value of 0.015 and F value of 3.966. and and the effect of physics is also statistically significant with a p value of 4.186E-16 and F value of 3.966. However, the interaction between two is not statistically significant.
A similar two-way ANOVA was also performed for cube manipulations per minute with/without Oculus and with/without Shadows. The results indicate that the effect of Oculus is statistically significant with a p value of 0.002 and F value of 3.966. But the effect of shadows and the interaction between the two conditions is not statistically significant. Average number of cube manipulations per minute are as shown in Table 1 and Table 2. It can be summarized from these tables that the users performed better overall without the Oculus than with the Oculus. They performed significantly well with physics enabled than without it. The presence of shadows only slightly benefited the users' performance.

With Oculus	
With shadows	0.346 cubes/min
Without Shadows	0.275 cubes/min
With Physics	0.402 cubes/min
Without Physics	0.051 cubes/min

Table 1: Performance analysis for tasks with Oculus

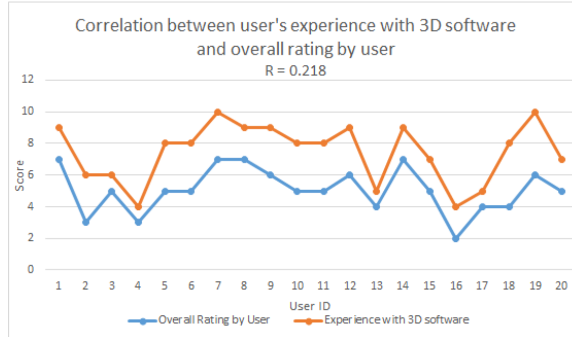

Figure 7: Correlation between user's experience with 3D software and overall rating by user R = 0.218

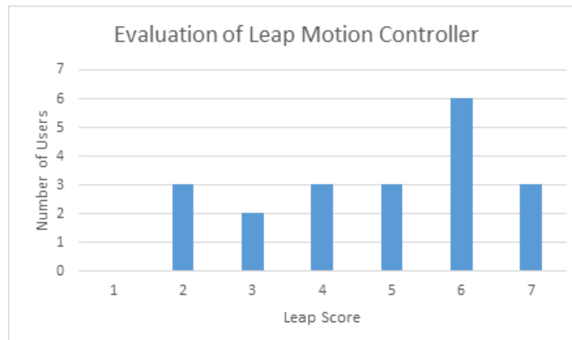

Figure 8: Evaluation of Leap Motion Controller

Without Oculus	
With shadows	0.503 cubes/min
Without Shadows	0.42 cubes/min
With Physics	0.564 cubes/min
Without Physics	0.087 cubes/min

Table 2: Performance analysis for tasks without Oculus

5.2 Qualitative results

The proposed system was ranked high by the users for the factors of ease of assembly and gestural information recognition. They expressed that the rendering of their hands in the environment greatly increased the sense of presence and general awareness of their actions.The choice of intuitive gestures such as pinching helped the users to interact with the system with a very gentle learning curve. While most of the users found it easy to interact with the virtual objects, very few users found it hard to initiate the interaction with the virtual objects for ease of interaction. This was due the inability of the Leap Motion sensor to track the fingers of the users outside of its optimal tracking region.
We took an interactive feedback from the users about how useful would they find our system design for their applications. Many users connected the system immediately to their backgrounds. One of the participants, a mechanical engineer expressed how this system would be help in building collaborative 3D designs of fluid flows with a remote team, often the clients. Their current method includes the use of video conferencing system or remotely logging into the remote teams computer. A more extensive version of our proposed system could allow both teams to actively participate and visualize the end product in an enhanced way.

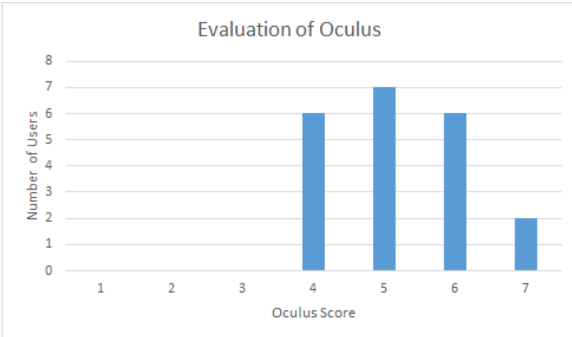

Figure 9: Evaluation of Oculus

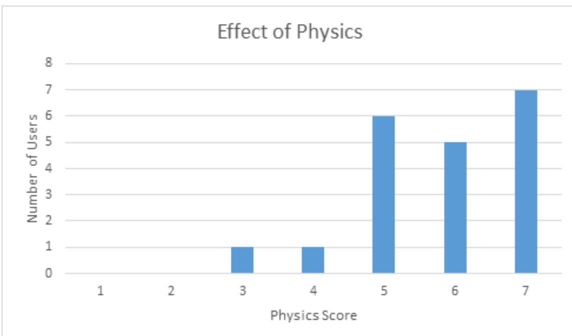

Figure 10: Evaluation of presence of physics

6 Conclusions and Future Work

We demonstrated and evaluated a 3D User interface platform to facilitate collaborative object arrangement in a Virtual environment. Our system takes advantage of the most recent advances in display and sensing technology, combining the Oculus Rift and finger tracking through the Leap Motion to merge the physical space with virtual objects, which users can manipulate using simple hand gestures. The initial experimentation was also done using Kinect as the hand gesture recognition tool but the large space requirement of the Kinect made Leap Motion a better choice to have better use of limited physical space. The qualitative results showed overall user satisfaction and ease of handling and working with the system. The quantitative results exhibited positive correlation between PURDUE SPATIAL VISUALIZATION test [4] score and overall performance which implies that user with good 3D spatial sense performed better in assembling the target model. The other observation was that gender bias did not affect the user performance but experience with 3-D software, gaming does seem to correlate with superior user performance.

We hypothesized that the higher-fidelity display and cues would improve objective and subjective performance. This hypothesis was validated since the variation of the presence of Physics and shadows helped the user to perform the task efficiently. The variation of the third parameter which was the display technology also affected the performance as the Oculus did pose additional overhead on the performance of many users. This clearly indicated that some learning curve is associated with using HMDs. We also successfully answered the research questions posed at the beginning of the study, such as the: effect of variation of factors, scalability of the application, users ease and performance metrics. Thus, the proposed system provided a seamless interface to facilitate users to model their assembly designs using 3D technology and also successfully incorporates collaborative, remote user interaction.

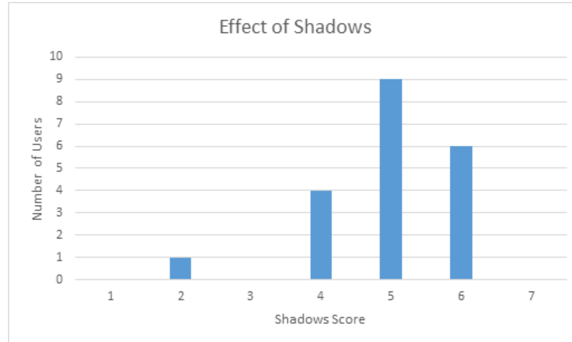

Figure 11: Evaluation of effect of shadows

Improving object manipulation holds immense potential to enhance current system design. Challenges are abundant in this domain since the general population is not yet accustomed to truly natural, 3D user interfaces, especially when there is no haptic feedback. Therefore, introducing haptic feedback for the current VR-based interface is a good direction to explore. In addition, setting up an interaction area and integrating local and wide field tracking could make the system more robust to different scenarios. This would involve the design of an interaction space where the physical location of a person can be tracked in real world. The addition of a haptic or sound feedback to user on drop, collision or interference will also give an enhanced immersive experience to the user. In the existing implementation, the hands might be lost by the tracking system or the system may return unreliable finger joint positions. This acts as an obstacle as this severely limits the effective interaction range of the system to the Leap Motion sensors optimal tracking region. A better choice for the sensor, a high-resolution depth camera, would help track the hands in a larger region of space.

Acknowledgments

The authors wish to thank ICAT, Dr. Benjamin Knapp, Visual Computing Lab at Virginia Tech. This work was partially supported by NSF award number 1305231.

References

[1] P. Apostolellis, B. Bortz, M. Peng, N. Polys, and A. Hoegh. Poster: Exploring the integrality and separability of the leap motion controller for direct manipulation 3d interaction. In *3D User Interfaces (3DUI), 2014 IEEE Symposium on*, pages 153–154, March 2014.

[2] D. A. Bowman and R. P. McMahan. Virtual reality: How much immersion is enough? *Computer*, 40(7):36–43, July 2007.

[3] J. C. Coelho and F. J. Verbeek. Pointing task evaluation of leap motion controller in 3d virtual environment. In E. Fiume, editor, *Proceedings of SIGGRAPH 2001*. van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR, 2014.

[4] R. B. Guay. Purdue spatial visualization tests. https://lphscadd.wikispaces.com/file/view/Purdue+Spatial+Test.pdf.

[5] D. Kartch. *Efficient Rendering and Compression for Full-Parallax Computer-Generated Holographic Stereograms*. PhD thesis, Cornell University, 2000.

[6] H. Landis. Global illumination in production. ACM SIGGRAPH 2002 Course #16 Notes, July 2002.

[7] F. I. Parke and K. Waters. *Computer Facial Animation*. A. K. Peters, 1996.

[8] F. Pellacini, K. Vidimče, A. Lefohn, A. Mohr, M. Leone, and J. Warren. Lpics: a hybrid hardware-accelerated relighting engine for computer cinematography. *ACM Transactions on Graphics*, 24(3):464–470, Aug. 2005.

[9] O. H. D. K. S. I. M. W. A. D. Wilson. Holodesk: Direct 3d interactions with a situated see-through display. 2012.

[10] Y. L. H. Yee. Spatiotemporal sensistivity and visual attention for efficient rendering of dynamic environments. Master's thesis, Cornell University, 2000.