

Evaluating the Effect of Positional Head-Tracking on Task Performance in 3D Modeling User Interfaces

Max Krichenbauer^a, Goshiro Yamamoto^a, Takafumi Taketomi^a, Christian Sandor^a, Hirokazu Kato^a, Steven Feiner^b

^aNara Institute of Science and Technology

^bColumbia University

Abstract

With the ongoing introduction of wide-FOV VR head-worn displays into the consumer market, the application of VR 3D UIs to professional work environments is attracting increasing attention. One of the most conspicuous concepts is immersive 3D modeling and content creation. In spite of the long research history and multitude of proposed systems, there have been very few analyses of the effect of 3D UIs on productivity in 3D design.

In this work, we explore the effect of positional head-tracking and its accompanying parallax depth cues on task performance in 3D object selection and transformation in a setting modeled after 3D design work. Previous studies have come to different conclusions on the importance of positional head-tracking and did not investigate professional 3D modeling tools. In contrast, we performed a user study with design students using professional software on a task that closely emulates their work. Surprisingly, we did not find a significant effect of positional head-tracking on task-completion time, neither when using a traditional 2D mouse nor when using a pinch glove as a 3D input device. Furthermore, we found that the users worked significantly faster with the mouse. We discuss possible explanations and implications for the design of 3D UIs.

Keywords: Augmented Reality, Computer Aided Design, 3D User Interfaces

2010 MSC: 00-01, 99-00

1. Introduction

3D design is one of the oldest and most frequently suggested uses for immersive 3D user interfaces (UIs), dating back to the 1970s [1]. Since then, many prototype 3D UIs have been suggested and at least partially evaluated, some of which show great promise. However, there have been relatively few studies that analyze workflows typical of 3D design and evaluate 3D UI concepts that fit these characteristics.

In our previous research we analyzed the real-life work situation of 3D media professionals showed that 3D UIs or even 3D input devices have so far failed to appeal to designers for adoption in their daily work [2]. This may be related to a lack of understanding of specific UI design factors relative to the requirements of this field of work.

In this paper, we attempt to obtain a better understanding of how certain 3D UI concepts relate to the workflow commonly found in professional 3D design for games or movies, such as modeling, rigging, animation, simulation, and lighting/rendering. While professional design work includes some work steps that are inherently 2D (such as compositing or color-grading), we ignore these tasks and focus on those work steps that are 3D. We furthermore focus on 3D object selection and transformation, since they are of great relevance to many real-life 3D design tasks.

In a preliminary pilot study to explore how artists use their current 2D mouse-and-keyboard UIs (see Section 3), we noticed a particular recurring behavior of “wiggling the view-

point,” wherein the artist rapidly and repeatedly changes the position of the virtual camera by small amounts, apparently in an attempt to gain a better understanding of the 3D shape of the virtual scene in the editor by emulating head-motion parallax.

From this, we theorized that a full-fledged immersive 3D UI could offer increased efficiency, at least in part because positional head-tracking (also called head-coupling or viewpoint-dependent imaging) would provide the necessary parallax effect automatically and probably even without conscious action by the artist.

Previous studies of the effect of positional head-tracking on user performance in various task settings have come to contradictory conclusions: some have found it beneficial [3, 4, 5, 6], some have found no effect [7, 8], and some have even found a negative effect [3, 9]. However, none of these addressed the challenges of 3D design. Since different tasks impose different requirements on the UI, studies performed with a generalized task and participants chosen from the general population may not apply to a particular real-life work environment. To make our results relevant to 3D design, we conducted a user study with 3D artists on a 3D object selection-and-transformation task and UI that are based on 3D design work (see Figures 1 and 2). Our within-subject experiment examined performance while wearing an orientation-tracked stereoscopic head-worn display (HWD), comparing a 2D mouse with a 3D pinch-glove, and the presence or absence of positional head-tracking.

Our results show no statistically significant effect of positional head-tracking on task performance, neither when using

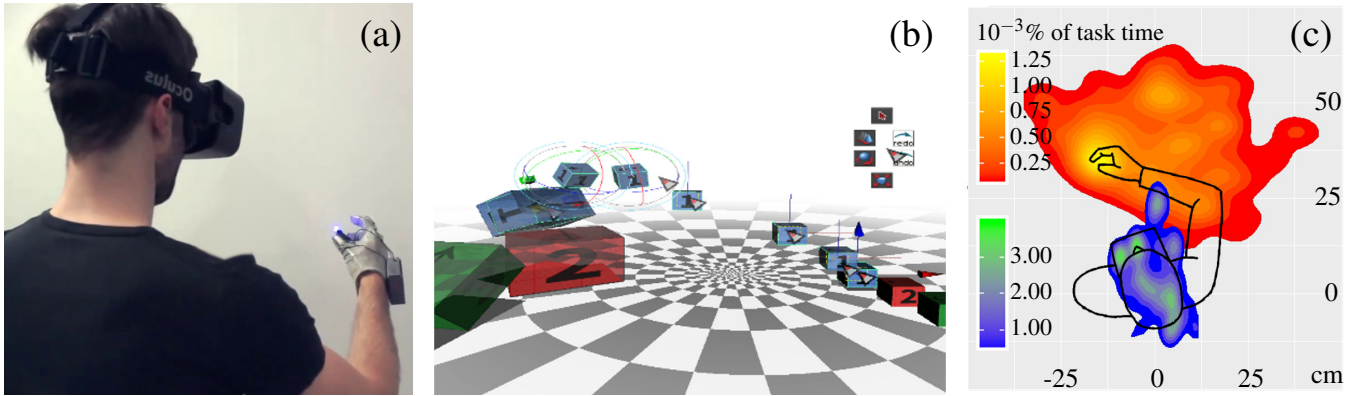


Figure 1: Our study analyzed the effect of positional head-tracking on task performance in a 3D object placement task that was modeled closely after 3D design work and was performed by experienced designers. (a) The experimental set-up in the glove condition. (b) Illustration of the task: Translating, rotating, and scaling objects on the right to match objects on the left. Stages in the transformation of the blue object are shown in this composite. (c) Heat-map indicating the head (blue) and hand (red) motion of one participant (top view).

the 2D mouse nor when using the 3D glove to perform a 3D object selection and transformation (translation, rotation, and scaling) task based on the creative design workflow. Furthermore, participants were able to perform the task significantly faster when using the mouse instead of the glove.

Our observations point to the possibility that stereo vision provided the necessary depth cues, making participants somewhat reluctant to move their heads at all when wearing the HWD, and that using the glove created challenges for the artists, who were already familiar with the mouse-based UI.

Our work thus makes the following contributions: Our results question the common belief in the importance of parallax depth cues from positional head-tracking. While positional head-tracking can increase the sense of immersion and may be crucial for some tasks, some 3D design-related tasks may not profit significantly. This further points to the possibility that VR work environments for 3D design can be created more easily and inexpensively by omitting hardware required for positional head-tracking (which is commonly achieved either by additional environment-mounted tracking hardware or computationally expensive visual odometry algorithms). We provide this insight in relation to both a 2D input device and a 3D input device. Based on empirical findings and qualitative feedback, we expand the knowledge base for 3D UI designers regarding practical implementation. Our work also raises new questions regarding the individual and combined effects of UI principles on task performance in 3D design work.

2. Related Work

Previous publications related to our efforts towards immersive 3D UIs for professional media production can be classified into two categories: analysis of positional head-tracking as a human factor in 3D UIs, and suggested prototype UIs for 3D modeling.

Studies on positional head-tracking in 3D UIs. Several publications have analyzed the effects of positional head-tracking on task performance in VR and AR 3D UIs, but came to contradictory conclusions. See Table 1 for an overview of

the publications we discuss and their respective results and limitations.

Arthur et al. [3] analyzed the effect of stereo vision and positional head-tracking on task performance in fish-tank (i.e., monitor-based) VR systems with a non-interactive cognitive task. Participants were shown two intertwined 3D tree structures and were asked to assign a given leaf to either of the two root points. The study compared task completion time and error rates for several conditions, including a monoscopic 3D image, a stereoscopic 3D image, a monoscopic 3D image with correct (head-coupled) perspective, and a stereoscopic 3D image with correct (head-coupled) perspective. Participants were instructed to move their heads and try to make as few errors as possible rather than optimizing time performance. They found that head-coupling alone (without-stereo) was slower than with stereo and even slower than static monoscopic and stereoscopic images, but that it decreased the error rate. In the stereo viewing condition, however, they found that positional head-tracking improved both time and error rate.

Similarly, Ware and Franck [6] performed experiments to determine which depth cues help participants perceive complex 3D graphs correctly in a fish-tank VR environment. They found that binocular stereo 3D improves performance by a factor of 1.6, parallax motion from positional head-tracking by 2.2, and a combination of both by a factor of 3. They consistently found that motion parallax has a stronger effect than stereo 3D, but recognized that the source of the viewpoint motion does not have to stem from head motion. Mouse input or a predefined slow rotational motion appear to work just as well.

These results provide some evidence for the effect of positional head-tracking, but are not directly applicable to 3D design work performance because the task was non-interactive and observational.

Boritz and Booth [7] also published a fish-tank VR study for a 3DOF (translational) target-pointing task. In this study, they found that binocular stereo allowed significantly faster performance than monoscopic imagery, but could not find a sustainable effect from positional head-tracking. During the first trials, positional head-tracking significantly improved performance in

the monoscopic display condition but degraded performance in the binocular stereo condition. However, this effect quickly wore off as participants adjusted to the task. A second study by the same authors analyzed a 6DOF placement (object-docking) task, but positional head-tracking was not among the conditions tested [11].

Teather et al. [8] performed another fish-tank VR user study in which participants performed basic translational positioning tasks under different viewing and interaction conditions, including positionally head-tracked stereo 3D. They found that stereo 3D has a positive effect on error rates (though not on task completion time), but could not find evidence for any effect from positional head-tracking. While these results give evidence that positional head-tracking is not beneficial to 3D interaction task performance, they too are not directly applicable to 3D design work because the task was intentionally chosen to be simplistic. The authors focus on novice users and limited the interaction to selection and surface-registered translation, ignoring the complexity of object rotation and scaling, as well as the possibility that objects may float in space, thus making the task essentially 2D.

One issue that these studies have in common is that they used fish-tank VR systems, for which the possible range of head motion is limited with respect to the position of the virtual object relative to the viewer. It is not possible to “look around” the virtual object or alter the perspective significantly. Furthermore, in these studies, there was a local separation between the input device and the virtual objects. The participants would interact with the input device next to the screen where they would see the result, which might have affected the outcome.

Arsenault and Ware [4] performed a user study whose interaction took place behind a mirror in which the participant sees the virtual environment, thus achieving perfect alignment of real hand and virtual 3D cursor. Their results show significant improvements when positional head-tracking is enabled. However, in their experimental design, they force participants to change their viewing angle by about 18°, causing a misalignment of the hand and virtual cursor when positional head-tracking is disabled. So the effect may have resulted from this misalignment, since users might not have naturally shifted their head position that far.

Bajer et al. [9] tested a selection task in a fish-tank AR system in which the participant was close to the screen and the 3D input device was aligned with the virtual pointer. They compared task performance with a control condition that used a 2D mouse. Their results show that positional head-tracking actually made participants slower when the perceived height difference of selection targets increased. However, this may have resulted from participants spending more time to move the pointing device “upwards” instead of just from side to side, indicating increased depth perception, which in this case happens to be detrimental to task performance.

Sandor et al. [5] performed a user study on object selection performance using a haptic device and a video see-through (VST) HWD. In half their conditions, they simulated a half-mirror-based VR system, displaying a virtual semi-transparent screen floating over the work area on which the virtual ob-

Table 1: Overview of user studies on the effect of positional head-tracking, grouped by their result. The right column mentions limitations in the respective study design.

Positional Head-Tracking Improves Performance:	
Arthur et al. [3]	Only stereoscopic passive viewing task. Fish tank VR.
Arsenault and Ware [4]	Forced people to move their head (causing misalignment without positional head-tracking).
Sandor et al. [5]	Simulated half-mirror (includes other effects such as brightness).
Ware and Franck [6]	Only passive viewing task. Fish-tank VR.
No Effect from Positional Head-Tracking:	
Boritz and Booth [7]	Fish-tank VR.
Teather et al. [8]	Fish-tank VR. Focused on novice users.
Jones et al. [10]	Only tested depth-estimation ability.
Positional Head-Tracking Decreases Performance:	
Arthur et al. [3]	Only monoscopic passive viewing task. Fish tank VR. Decreased speed, but lowered error rate.
Bajer et al. [9]	Fish-tank VR. Effect possibly caused by 3D motion, not “on-screen” motion.

jects were visible. This condition (among other differences) provided no positional head-tracking and was found to significantly decrease performance. However, they did not explicitly investigate the effect of positional head-tracking and only mentioned it as one of three possible explanations for the observed effect. Furthermore, there was no virtual cursor displayed in the simulated half-mirror condition, which meant that any head motion from the center of the virtual camera would make it impossible for the participant to know precisely where they were pointing at in the virtual scene.

Using an HWD, Jones et al. [10] measured participants’ ability to estimate the distance of objects in real, VR, and AR conditions, both with and without cues from head-motion parallax. Their study showed no benefit from lateral head motion for virtual objects, and only negative effects for real objects when wearing an inactive optical see-through (OST) HWD, whose only intended effect was to add artificial inertia to the participants’ head motions. However, objects in this study were positioned several meters away from the participant and not within arm’s reach, as they would be in a workplace environment.

In addition, none of these prior studies on the effect of positional head-tracking used current-generation HWDs, which may have affected task performance due to the increased resolution, improved tracking performance, a larger field of view, increased freedom of movement, and a greater sense of immersion that modern devices offer over previous systems. They also considered only fairly generic tasks and recruited either subjects from the general population or engineering students with little or no experience in 3D modeling; thus, their results might not directly apply to 3D design.

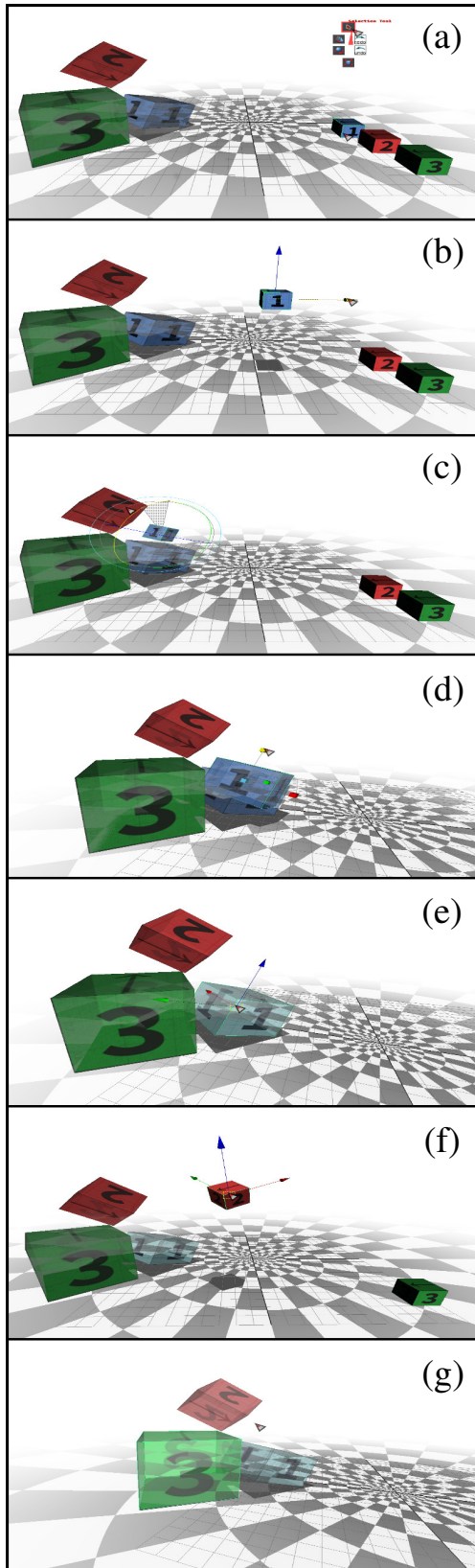


Figure 2: Study task as seen through the HWD. (a) The initial 3D scene. Source objects are on the right, and transformation goals are on the left. The tool menu appears at the top. (b) Translating the first object (blue cuboid labeled “1”). (c) Rotating the object. (d) Scaling the object anisotropically. (e) After the first object is correctly transformed, it is highlighted, and the goal disappears. (f) Transforming the second object. (g) All three objects transformed correctly.

Prototype systems of 3D UIs for 3D design. Given the long history of research on 3D UIs, a multitude of prototype systems have been developed and sometimes formally evaluated.

LaViola and Keefe [12] published a review of 3D UIs for art and design. Fiala developed ARpm [13], an AR front-end to Autodesk 3D Studio Max. However, its applicability to real-life design work is limited because it functions by taking screenshots of the software and sending Windows system calls to control it, severely constraining performance and usability.

Takala et al. [14] implemented a semi-immersive 3D UI for the Blender modeling software using PlayStation Move controllers. They tested their prototype in a pilot user study with seven professional 3D artists but limited their evaluation to whether it was fun to use or whether the participants experienced fatigue. Interestingly, almost all of the artists agreed with the expectation that 3D UIs will become commonplace for 3D design within this decade.

Other examples are the entries to the 2013 IEEE 3DUI Contest. Two of these are the DIY World Builder [15] and the Wonderland Builder [16], which are aimed at level creation for games or creating fantasy worlds.

Jackson and Keefe [17] present a different approach where the artist can start modeling from scratch by sketching lines and shapes in 2D outside of VR and then importing these into a CAVE system where they are “lifted off” of a plane and interactively positioned and edited in 3D space.

Finally, Mine et al. [18] developed an immersive VR modeling application based on SketchUp with the intention to enable professional adoption like commercial-off-the-shelf programs. Their system combines two hand-held smart-phones with additional appendages with one of four different VR systems — none of which is an HWD. While they occasionally invited in 3D artists to evaluate their approach, the authors themselves served as the primary users of their system. Through their work on the prototype system, they derive nine interaction design guidelines, including using 2D touch devices for precise input and using haptic feedback.

3. User Study

We began by working with artists to understand their work environment. In August and September 2015, we performed a pilot study with one professional artist and one amateur, in which we examined their habitual 2D UI workflow. We instructed them to continue working on their own projects as they normally do, while we recorded their actions with a video camera pointed at their computer. We left the room while recording, to avoid influencing the participants’ behavior. The participants worked on their own computers which featured a mouse and keyboard as the only input devices. Both of the participants engaged in the editing of polygonal models.

Analyzing their 2D workflow, we found several interesting actions in which both the amateur and the professional substantially engaged. One was the frequent and rapid change of viewpoint (camera position) on a very small scale. This camera motion did not allow the artist to see a different object or a previously hidden side of the object, and often ended very close to

the original position. This behavior—where the artist returns to the original viewpoint and continues working on the same part of the object or scene as before—makes up about 42% of all viewpoint changes (3D scene navigation) and about 8% of total work time on average. In comparison, the artists spent between 4% (Professional) and 15% (Amateur) of the work time looking at reference material, and between 47% (Professional) and 49% (Amateur) on 3D object transformation. The intended purpose of these minute viewpoint changes seems to be to gain a better spatial understanding of the virtual 3D object. Since the monitor is monoscopic and no parallax effect from head motion is available, the visible image is ambiguous in its depth and 3D shape. A slight “wiggle” of the virtual camera produces a parallax motion that gives the artist a better understanding of the virtual scene. This finding is in accordance with our previous findings [2], in which 3D artists reported in a survey that they used the camera controls of their 3D software extensively or even constantly.

Therefore, we theorized that in a VR work environment, positional head-tracking might make this operation unnecessary, possibly improving work performance. To validate this conception we formulated the following hypothesis:

H_1 : Enabling or disabling positional head-tracking has a significant impact on task performance in a 3D selection and transformation task. Assuming that positional head-tracking in an immersive VR environment would provide these cues sufficiently to make wiggling the camera unnecessary, we hypothesized that disabling positional head-tracking would negatively impact performance. If this hypothesis would be falsified, it would allow for VR work environments that are less expensive and easier to set up because the rotational information from an internal inertial measurement unit (IMU) in the HWD is sufficient, and no additional hardware for positional tracking is required.

Some prior studies have found that depth cues can have different or even opposite effects based on whether a 2D input device or 3D input device was used [9]. Therefore, we decided to test our hypothesis with both a traditional mouse and a 3D input pinch-glove.

3.1. Experimental Platform

To test this hypothesis, we developed a prototype 3D modeling UI. In order to stay as closely related to our target field of application of 3D design, we based our UI on Autodesk Maya, by developing our UI to be a plug-in for this software. The plug-in makes it possible to use Maya with an Oculus Rift DK2. Our test system consisted of a Dell Precision Notebook with a Intel Core i5 CPU with 2.90 GHz, 8GB RAM, and an nVidia Quadro K4100M graphics adapter. The computer was running Microsoft Windows 7 and Autodesk Maya 2014.

We performed the study with Oculus SDK 0.6.0, with dynamic pose prediction based on internally measured latency enabled. When we later updated the SDK to 0.8.0, we measured an average Motion-to-Photon Latency of 16ms.

The input device was either a normal 2D mouse or a 3D pinch glove. The mouse was a wired laser mouse (a Dell MOC5UO),

which is the most common input device in 3D design work [2]. When the mouse was used, a 2D cursor was displayed in the dominant-eye view only and thus was parallax free (similar to looking through a red-dot sight or reflector sight of a rifle). The mouse cursor was projected on a virtual plane that was at 10cm distance and perpendicular to the HWD. It always followed the HWD’s motion, so a head motion would not result in a visual motion of the mouse cursor as long as the mouse remained stationary. Since the cursor was displayed to a single eye, it did not actually appear to be at a specific distance. Selection was performed by ray-casting from the dominant eye. The UI was the same as in Autodesk Maya with two exceptions: a marking-menu [19] on the right mouse button to select the tool (translation, rotation, and scaling), and viewpoint navigation on the middle mouse button. When the middle mouse button was pressed, the mouse controlled the viewpoint to support tumbling around the selected object. The mouse wheel allowed dollying forward and backward to focus on certain areas of interaction. This is similar to the camera motion used in Maya and other professional modeling software products, so we assumed it would be immediately understandable to the participating artists.

We constructed the 3D input glove from a thin cotton glove in which we sewed conductive threads for the pinch buttons. Prior to using our own glove system, we tried a LeapMotion Controller¹, but found the hand tracking to be too imprecise to detect pinch gestures with different fingers reliably. Two different sizes of the glove were created in both left-hand and right-hand versions, to accommodate differences in hand size and handedness. Each glove featured up to eight buttons, of which only the most basic four were used in this study. These buttons were the main interaction button (similar to the left mouse button), a button to invoke the tool menu (similar to the right mouse button), an undo button, and a navigation button, which allowed changing the virtual viewpoint with a “grabbing-the-air navigation” metaphor [20]. An Arduino microcontroller and Bluetooth modem were attached to the glove to transfer pinch actions wirelessly to the computer.

A NaturalPoint OptiTrack Flex3 motion-capture camera system was used to track the position of the thumb, which acted as the 3D cursor. We chose the thumb as the frame of reference since it is more positionally stable during a natural pinch gesture than the finger. This also improved the tracking quality and comfort since we did not have to track every finger individually. While the Maya UI usually requires the users to interact with the displayed manipulator handle of the tool (e.g., in order to select the appropriate axis of translation), we removed this requirement in the glove condition. The reasons for this were that it is unnecessary (since the 3D motion of the hand provides the correct axis of interaction) and it made the UI hard to use (because pixel-precision interaction in free space is almost impossible to perform reliably). Our set-up featured six OptiTrack cameras: four in front of the participant (two facing down from above and two facing up from below) and two behind the participant (viewing the work area over the participant’s shoulder).

¹www.leapmotion.com

The total area of the tracked space was approximately $1.5 \times 1.5 \times 1.5$ meters.

The users were seated on an armless chair at the far end of this tracking volume, looking into its center. In the glove condition, the participants had no possibility to rest their elbows and we did not observe participants bending over to rest their elbows on their knees. However, it was possible to perform the task with bent arms which they naturally did. In the mouse condition, a small table was positioned beside the chair so that the participants could rest their mouse hand on the table.

We used Tsai's hand-eye-calibration algorithm [21] to align the coordinate systems of the Oculus Rift DK2 and OptiTrack. OptiTrack software was run on a separate networked computer, to ensure good performance. When using the glove, a 3D arrow was rendered in stereo in the virtual environment at the location of the participant's thumb. The UI was the same for both mouse and glove conditions, except for the addition of a "6DOF Tool" for the glove, which allowed simultaneous control of translation and rotation. The control/display ratio was 1:1 in the glove condition, and 1cm:14.7° FOV in the mouse condition. The total FOV of the Oculus Rift DK2 was about 106° vertically and 95° horizontally.

In the normal state, the system presented a fully immersive 360° virtual environment with both positional (lateral) and rotational head-tracking. Head-tracking was performed by the Oculus Rift which was very stable and free of noticeable lag. Positional head-tracking could be turned off and on by the experimenter at any time. When positional head-tracking was turned off, it produced the effect of looking at a virtual 3D monitor large enough to fill the complete FOV when looking straight ahead, and following all translational head motion. Further, without positional head-tracking, the correct alignment between the real hand and 3D cursor would be broken if the user's head moved. For example, if the user were to move their head sideways, the virtual screen would move with it to display the 3D cursor at the new position, even though the hand had remained stationary, thus resulting in a translational offset between the real hand and cursor. It should be mentioned that in a completely virtual environment, such an offset would not necessarily cause confusion, just like the mouse being in a different place than the on-screen mouse pointer would not confuse 2D desktop software users. Independent of positional head-tracking, rotational head-tracking was always enabled, so the user did not have to rely on peripheral vision to view any part of the work area. This is important because the Oculus Rift DK2 exhibits significantly more distortion toward the periphery of each panel.

Since we are using the same prototype for developing AR UIs for more complex tasks, it has additional features such as front-facing cameras that turn the Oculus Rift DK2 into a VST HWD, and additional UI functionality to access additional functionality in Maya. These features were not available to the participants during the study tasks they performed; however, participants were allowed to use the whole prototype freely for 30 minutes prior to performing the tasks. This was done in order to evaluate the whole AR-enabled prototype, which is beyond the scope of this publication. However, it also guaranteed that

all participants had at least basic experience in working in an immersive modeling system and knew the basic UI used for the task performance trials.

We decided to perform the user study on positional head-tracking in VR instead of AR, because AR without positional head-tracking can be quite confusing (giving the sensation of the whole room moving with you whenever you move your head). Therefore, we did not use the cameras in all conditions and instead presented a simple virtual work area consisting of a circular base.

3.2. Procedure

We conducted a user study in December 2015 at Kyoto Saga University of Arts with nine participants: seven design students, one faculty staff member, and one professional 3D artist (five female, ages 19–35, mean age 22.2). All had at least one year prior experience with 3D design software.

In order to get a better understanding of the level of expertise of our participants with 3D design, we asked them to fill out a form on their previous experience. For a time frame of the past ten years, we asked the participant to estimate the number of hours spend each week working with 3D design software, averaged over the whole year. So if a participant worked on a project for an average time of 3 hours for half a year, then the average time spent over the whole year would be 1.5 hours per week. From this, we calculated an estimate of the total hours of experience by multiplying each annual estimate by 50 weeks per year (assuming that participants tend to neglect holidays). Our participants reported experience levels between 75 hours (least experienced participant) and 10000 hours (most experienced participant). The average experience level was around 1200 hours. While this is naturally an extremely coarse estimation, it still gives us some information about the likely level of expertise of the participant on a logarithmic scale. Users whose experience is limited to tens of hours can be assumed to be beginners (two participants), those with hundreds of hours to be amateurs (six participants), those with thousands of hours to have achieved expert level (one participant). This is in accordance with prior findings from different fields of artistic endeavor [22].

Before the participants started using the prototype, we determined their ocular dominance with a Miles test (five right-eye dominant), asked them about their dominant hand (seven right-hand dominant), and informed them about the risk of cybersickness when using an HWD. They then tested the glove-based 3D UI for 30 minutes, following a tutorial. This allowed them to get used to immersive modeling and learn how to use the glove. During this period, we switched the prototype to AR-mode, streaming video images from a pair of cameras attached to the front of the HWD, turning it into a VST HWD. This reduced disorientation and allowed the experimenter to point things out to the participant in order to help them. After this session, a rest period was given, during which the participant filled out some forms. The cameras were then turned off and remained off during the entirety of the timed study trials.

Then, participants were timed on a 3D object selection-and-transformation task with simple 3D objects. Two groups of 3D

boxes were displayed, one set being the “goal” arrangement, the other being the “source” objects to place. The task was to transform each “source” object in the same way and to the same place as its corresponding “goal” object. This task involved nine degrees of freedom (DOF): three DOF each for translation, rotation, and scaling. The “goal” arrangement was chosen randomly from a set of ten manually prepared scenes in which the objects formed a pile. Each of these ten arrangements was very similar in that one cube was resting on the ground, one leaned against it, and one was balanced on top of the others; however, the exact place, orientation, scale, and order of cubes was different in order to avoid learning effects. The “source” boxes always started on the right side of the work area and were uniformly placed and scaled and aligned side-by-side with each other.

The task was designed to emulate artistic 3D modeling, in which the artist starts with a certain goal in mind (either provided by a concept artist or art director, or of the artist’s own imagination) and tries to reach this goal. We have chosen this task for our analysis because 3D selection and transformation of whole objects—together with tool selection which was also an element in our UI—is common in 3D design work, but also because it approximates a large portion of more complex 3D interaction tasks. For example, skeletal animation can be seen as the selection and manipulation of “joint” objects and animation handles, fluid simulation as creation, selection, and manipulation of “emitter” and “effector” objects, and sculpting as manipulation of a “virtual sculpting tool”. We designed our task based on observations from our video analysis of two artists using the traditional 2D UI, in which they spent almost half of their time on object selection and 3D transformation. Of course, purely observational tasks may therefore not be represented correctly in our study.

We always grouped three 3D objects at a time, since artistic scenes are rooted in the relative arrangement of objects—an isolated single-object transformation task would be less representative of the tasks we are targeting. The work area was 70cm wide from side to side, initially displayed at a distance of 60cm in front of the participant and 35cm below eye level. Thus, it was possible to observe the whole area without requiring extensive rotational head motion (the work area consumed about 60° of the HWD’s 95° horizontal FOV). However, participants were still able to manually change the position of the work area to gain a better view of certain details. In the starting location, the cubes were approximately 5cm x 5cm x 5cm in size. In the “goal” configuration their size ranged between 10cm x 10cm x 10cm and 17.5cm x 12.5cm x 12.5cm (after anisotropic scaling). See Figure 2 for an example of the task. The tolerances for successful completion were approximately 2.5cm Euclidean distance, 5° of rotation around any axis (yaw, tilt, or roll), and a difference in scale of 5cm along any axis (x, y, or z). All three criteria had to be met for the object to be marked as correctly transformed.

Prior to starting the trials, the task was first demonstrated to the participant by having the experimenter briefly take over the UI from the outside and perform the transformation while the participant was wearing the HWD. This ensured that there

was no confusion as to the goal. During this demonstration positional head-tracking was enabled. Each participant was informed that time was the critical factor in task completion, but was also told not to work any more quickly than they found reasonable to perform the task.

For each trial, the participant used either the 3D glove input device introduced in the first (practice) part of the user study or the 2D mouse. While the task was the same in both conditions, it would be possible to perform it with fewer, more complex steps when using the glove, since it allowed 6DOF interaction which made it possible to perform translation and rotating simultaneously. During half the trials, positional head-tracking was switched off by the experimenter. However, rotational head-tracking was always enabled. The participants were not told during which trials positional head-tracking was enabled.

This study design yielded four different conditions: using a 2D mouse with positional head-tracking enabled, using a 2D mouse with positional head-tracking disabled, using a 3D input glove with positional head-tracking enabled, and using a 3D input glove with positional head-tracking disabled. The conditions were presented in a randomized order, but always with either both mouse conditions or both glove conditions first, never switching back and forth between mouse and glove. Switching the input device required the experimental set-up to be changed slightly, and this served as a brief resting period for the participant.

The first set in every block (three object transformations, during which head-tracking was enabled) was treated as a training set and removed from the sample. Some measurements were lost, due to technical problems or difficulties in the time schedule not allowing all conditions to be tested. The final analysis contains 30 three-object sets in the mouse conditions (performed by eight of the nine participants) totaling 90 object transformations, plus 14 three-object sets in the glove conditions (performed by five of the nine participants), totaling 42 object transformations.

We did not ask participants to fill out a standardized questionnaire such as SUS [23] or PSSUQ [24], since these provide measurements that are only relevant in comparison to other user interfaces. Our focus was solely on measuring differences in task performance related to positional head-tracking.

3.3. Results

A summary of the recorded measurements can be seen in Figure 3(a). When using the mouse, mean task-completion time was 314s without positional head-tracking (SD=111s), and 290s with positional head-tracking (SD=97s). When using the 3D input glove, mean task-completion time was 431.4s without positional head-tracking (SD=125.6s), and 459s with positional head-tracking (SD=127.2s). Analysis of within-subject performance showed a significant difference (defined by an α of 0.05) in the task performance between using the 2D mouse and the 3D input glove (average improvement of 159.4s; $p < 0.0035$; Figure 3b), but no significant effect from positional head-tracking, neither for the 2D mouse nor for the 3D input glove. (Mouse: difference of means $\approx 25.4s$ (7.7%), $p > 0.1$, Figure 3(c);

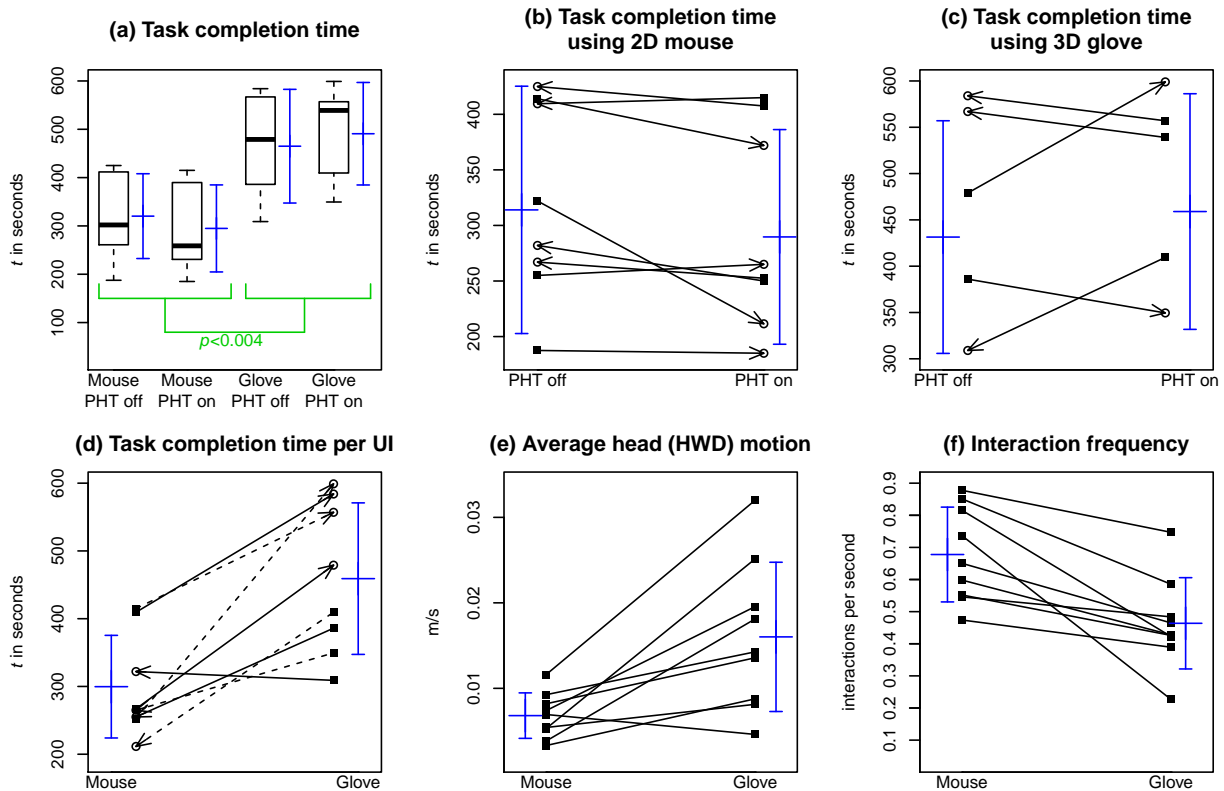


Figure 3: Experimental results. PHT indicates positional head tracking. Blue bars indicate mean and standard deviation. Data points are participant means. Dashed lines indicate measurements without positional head-tracking. Arrows indicate the order in which the conditions were performed.

Glove: difference of means $\approx -25.8s$ (-6.4%), $p > 0.49$, Figure 3d). We considered extending the user study to find more minute differences in task performance, but a power analysis using our sample to estimate population variance (i.e., assuming that future participants would exhibit a similar variability to previous ones) indicated that this was impractical, as we found that we would need $n > 77$ and $n > 76$ respectively for a test of power 0.95. This indicates that the expected effect of positional head-tracking is small compared to other factors. Our results show an effect size r of 0.1 (Cohen’s $d \approx 0.2$), which is considered small. Conversely, they express 95% confidence that the performance improvement for positional head-tracking is $< 17\%$.

We further analyzed the recorded motion data and found that participants moved their head significantly less when using the mouse ($p < 0.008$ on a within-subject t -test; see Figure 3e). Figure 1c shows an example of the recorded motion data of one participant as a heat-map. The motion volume in which the participants moved their heads was about $0.0066m^2$ on average when using the mouse, and $0.0257m^2$ on average when using the glove.

Another possible cause for the imbalance between mouse performance and 3D input-glove performance could be the familiarity with the traditional device, allowing for faster interaction. We therefore analyzed the recorded data, measuring the frequency of interactions (“clicks” for the mouse, and “pinches” for the glove) and found that the mouse was indeed

used more vigorously ($p < 0.0027$; see Figure 3f).

3.4. Discussion

Regarding our hypothesis, we have found no evidence for H_1 and therefore cannot confirm the assumption that positional head-tracking affects artist performance in a 9DOF object-transformation task, both for the 2D mouse condition as well as for the 3D input-glove condition. We will discuss possible explanations and implications in turn.

Not finding that positional head-tracking has a positive effect on task performance is in accordance with prior findings [10, 3, 8]. Their explanations for the lack of effect usually relate to the fact that while motion parallax can be an important depth cue, the natural range of motion of the head combined with the physical exertion make it an unattractive option to gain spatial understanding of the virtual object. The rather small amount of head motion observed in the participants supports this idea.

The difference in head motion between mouse use and glove use is likely to be the result of secondary motion. When using the 3D input glove, the hand has a wider range of motion in three dimensions, thus forcing the arm, shoulder, and subsequently the head, to move more to support hand interaction. The difference in head motion could have had an effect on task performance, but no such effect was observed. It is possible that the advantage of motion parallax was still too small in both cases to be noticed.

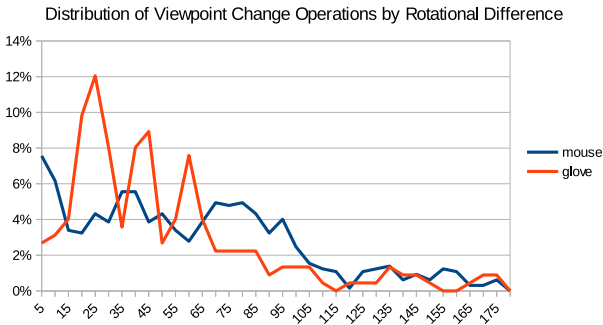


Figure 4: Frequency analysis of the viewpoint change operations by total rotational difference achieved through the operation.

One explanation for the limited range of motion may have been that participants felt hindered by the HWD. Three of the participants criticized the weight of the HWD after testing the prototype, giving respective estimations of the gravity of the problem of 4, 8, and 6 on a scale from 1 (negligible) to 10 (critical). It is possible that — burdened by the weight of the HWD — a head motion to change one’s point of view may actually be slower than a hand motion achieving the same effect, and thus not at all desirable by the participants.

Another possible explanation that is specific to our use case is that the “camera wiggle” has become so habitual that the user performs it even when the spatial relationships are clear. In our experimental platform, the users were able to use the middle mouse button or navigation button on the glove to change their viewpoint, in both positional head-tracking enabled and disabled conditions. While this feature was intended to allow major changes of the viewpoint necessary to perform the task, it may have been used by participants to perform the habitual “wiggling” even in the positional head-tracking enabled condition when a small head-motion could have sufficed. In order to test this assumption we analyzed the log files recorded during the experiment. Figure 4 shows a histogram with 5° bins of the viewpoint change operations performed by the user over the total rotational change achieved through the operation. Thus, the rightmost end of the graph (180°) means the user ended the operation looking at the objects from the opposite side of where the operation was started. While there was a slight tendency in the mouse condition for users to end close to where they started, this was not nearly as pronounced as in the 2D UI where “wiggling” made up 42% of all viewpoint change operations. Furthermore, no such tendency was present in the glove condition. Most viewpoint change operations fell into the range of 20° to 80°. The average was 50° in the mouse condition (SD: 43°) and 49° in the glove condition (SD: 36°).

An alternative interpretation for these results is that the moments of “wiggling” the camera in a 2D UI are used to consider the next steps. Thus, even when the wiggle is not necessary, the user might still pause the interaction regularly to consider what to do next, thus minimizing the possibility to improve task-completion time.

Finally, it is possible that the availability of stereo vision in

our UI provided enough spatial understanding to make the parallax motion unnecessary. This explanation, however, raises the question of why stereo glasses are not widely used in 3D design, since they would provide stereo 3D vision quite easily.

It is possible that positional head-tracking would prove more helpful in more cluttered scenes, where it may be necessary to move the head to gain a clear view of the area of interaction. However, in this case, the advantage of positional head-tracking would be artificial and not necessarily reflect the real-life work environment of artists accustomed to a 2D UI where clutter obstructing the line of sight is unacceptable.

Our results clearly show reduced performance when using a 3D input device, similar to previous work [8]. Given the great efficiency with which we interact with physical objects in our everyday life, it can be assumed that this stems, at least in part, from the artists’ familiarity with the mouse and maturity of the device. While the mouse has had a long time to mature into a reliable, precise, and universal interaction device, our gloves were research prototypes that all participants used for the first time ever. Further, the 6DOF tracking of the glove was not perfectly reliable, resulting in occasional jumping or misalignment of the virtual cursor. This is likely to have decreased performance in the glove conditions.

In fact, every single participant criticized the glove after testing the prototype UI. The estimated gravity of the reported problems on a scale from 1 (negligible) to 10 (critical) ranged from 1 to 8, with an average of 5.2. Common points of critique were the size and placement of the touch contact areas and the lack of haptic feedback.

It is also worth noting that the mouse was placed on a table, giving the participant’s hand more stability when using the system. This might have resulted in higher precision and in turn improved task performance. In addition, the pinching motion used with the 3D input glove can alter the 3D cursor position involuntarily. In our study, we could often observe participants struggling with finalizing the placement of the virtual object, because they were unable to end the interaction at the intended point.

4. Conclusions

Our results suggest that the effect of positional head-tracking on task performance in 3D design contexts, similar to those of our study, is not significant, at least in the short term. We have shown this for both a 2D mouse input and for a 3D input device. This suggests that future VR work environments for 3D design can be designed more efficiently by omitting positional head-tracking technology without significant loss in user performance. While absence of statistical proof does not imply complete absence of any effect, our statistical analysis gives both a rough estimate of the size of the expected effect as well as guidelines for the scale required of possible future experiments designed to determine the effect of positional head-tracking more precisely.

One weakness of the present study is that it does not account for long-term effects. It is possible that artists would get accustomed to using their head more, instead of their input device,

to change their viewpoint. In addition, deactivated positional head-tracking could have made the user more susceptible to cybersickness, thus reducing performance in the long run. In our user study, we asked participants after every condition whether they experienced cybersickness, but none reported more than a slight feeling of discomfort. However, since the question was intended only to ensure the participants' well-being and safety, we did not use a formal measurement tool. We also did not test the effect of positional head-tracking in the absence of stereoscopic 3D graphics, which may have compensated for the parallax depth cues under these conditions. Studying these effects is left to future work.

Acknowledgments

Steven Feiner was supported by NSF Grant IIS-1514429.

References

- [1] J. H. Clark, Designing surfaces in 3-D, *Communications of the ACM* 19 (8) (1976) 454–460.
- [2] M. Krichenbauer, G. Yamamoto, T. Taketomi, C. Sandor, H. Kato, Towards augmented reality user interfaces in 3D media production, in: *Mixed and Augmented Reality (ISMAR)*, 2014 IEEE International Symposium on, IEEE, 2014, pp. 23–28.
- [3] K. W. Arthur, K. S. Booth, C. Ware, Evaluating 3D task performance for fish tank virtual worlds, *ACM Transactions on Information Systems (TOIS)* 11 (3) (1993) 239–265.
- [4] R. Arsenault, C. Ware, Eye-hand co-ordination with force feedback, in: *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, ACM, 2000, pp. 408–414.
- [5] C. Sandor, S. Uchiyama, H. Yamamoto, Visuo-haptic systems: half-mirrors considered harmful, in: *EuroHaptics Conference, 2007 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics 2007. Second Joint*, IEEE, 2007, pp. 292–297.
- [6] C. Ware, G. Franck, Evaluating stereo and motion cues for visualizing information nets in three dimensions, *ACM Transactions on Graphics (TOG)* 15 (2) (1996) 121–140.
- [7] J. Boritz, K. S. Booth, A study of interactive 3D point location in a computer simulated virtual environment, in: *Proceedings of the ACM symposium on Virtual reality software and technology*, ACM, 1997, pp. 181–187.
- [8] R. J. Teather, W. Stuerzlinger, Guidelines for 3D positioning techniques, in: *Proceedings of the 2007 conference on Future Play*, ACM, 2007, pp. 61–68.
- [9] B. Bajer, R. J. Teather, W. Sturzlinger, Effects of stereo and head tracking in 3d selection tasks, in: *Proceedings of the 1st symposium on Spatial user interaction*, ACM, 2013, pp. 77–77.
- [10] J. A. Jones, J. E. Swan II, G. Singh, E. Kolstad, S. R. Ellis, The effects of virtual reality, augmented reality, and motion parallax on egocentric depth perception, in: *Proceedings of the 5th symposium on Applied perception in graphics and visualization*, ACM, 2008, pp. 9–14.
- [11] J. Boritz, K. S. Booth, A study of interactive 6 DOF docking in a computerised virtual environment, in: *Virtual Reality Annual International Symposium*, 1998. *Proceedings.*, IEEE 1998, IEEE, 1998, pp. 139–146.
- [12] J. J. LaViola, D. F. Keefe, 3D spatial interaction: applications for art, design, and science, in: *ACM SIGGRAPH 2011 Courses*, ACM, 2011, p. 1.
- [13] P. Fiala, N. Adamo-Villani, ARpm: an augmented reality interface for polygonal modeling, in: *Mixed and Augmented Reality*, 2005. *Proceedings. Fourth IEEE and ACM International Symposium on*, IEEE, 2005, pp. 196–197.
- [14] T. M. Takala, M. Makarainen, P. Hamalainen, Immersive 3D modeling with blender and off-the-shelf hardware, in: *3D User Interfaces (3DUI)*, 2013 IEEE Symposium on, IEEE, 2013, pp. 191–192.
- [15] J. Wang, O. Leach, R. W. Lindeman, DIY World Builder: an immersive level-editing system, in: *3D User Interfaces (3DUI)*, 2013 IEEE Symposium on, IEEE, 2013, pp. 195–196.
- [16] C. Barot, K. Carpentier, M. Collet, A. Cuella-Martin, V. Lanquempin, M. Muller, E. Pasquier, L. Picavet, A. Van Ceulen, K. Wagrez, The wonderland builder: Using storytelling to guide dream-like interaction, in: *3D User Interfaces (3DUI)*, 2013 IEEE Symposium on, IEEE, 2013, pp. 201–202.
- [17] B. Jackson, D. F. Keefe, Lift-off: Using reference imagery and freehand sketching to create 3d models in vr, *IEEE transactions on visualization and computer graphics* 22 (4) (2016) 1442–1451.
- [18] M. Mine, A. Yoganandan, D. Coffey, Principles, interactions and devices for real-world immersive modeling, *Computers & Graphics* 48 (2015) 84–98.
- [19] G. Kurtenbach, W. Buxton, User learning and performance with marking menus, in: *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, 1994, pp. 258–264.
- [20] D. P. Mapes, J. M. Moshell, A two-handed interface for object manipulation in virtual environments, *Presence: Teleoperators & Virtual Environments* 4 (4) (1995) 403–416.
- [21] R. Y. Tsai, R. K. Lenz, A new technique for fully autonomous and efficient 3D robotics hand/eye calibration, *Robotics and Automation, IEEE Transactions on* 5 (3) (1989) 345–358.
- [22] K. A. Ericsson, R. T. Krampe, C. Tesch-Römer, The role of deliberate practice in the acquisition of expert performance., *Psychological review* 100 (3) (1993) 363.
- [23] J. Brooke, et al., SUS-a quick and dirty usability scale, *Usability evaluation in industry* 189 (194) (1996) 4–7.
- [24] J. R. Lewis, Psychometric evaluation of the post-study system usability questionnaire: The PSSUQ, in: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 36, SAGE Publications, 1992, pp. 1259–1260.