Tampere University

Sriram Kumar Kishore Kumar

Designing hand-based interaction for precise and efficient object manipulation in Virtual Reality

Faculty of Information Technology and Communication Sciences (ITC)
Master's thesis
April 2021

ABSTRACT

Sriram Kumar Kishore Kumar: Designing hand-based interaction for precise and efficient object manipulation in Virtual Reality
Master's thesis
Tampere University
Master's Degree Programme in Human Technology Interaction
April 2021

Perceiving 3D anatomical data on a 2D screen is complicated, error-prone, and requires training for medical professionals as they mentally reconstruct 3D data from 2D. 3D data perceived in virtual reality reduces the 3D to 2D information loss, operating time, and cognitive load. Research in virtual reality applications for jaw osteotomy operation planning has explored interaction techniques using tracker, 3D pen, and haptic pen. However, medical professionals would like to use their hands as they do not need additional hardware and learn how to use it. Using hands as an input in virtual reality can be challenging because of the noisy hand tracking.In the process of jaw osteotomy operation planning, maximum accuracy is required for adjusting the osteotomy plane because this is the final step, and it compensates for the errors in the previous steps. This work focuses on designing and evaluating precise and efficient hand-based interaction techniques for plane alignment in virtual reality. A contextual inquiry is conducted to understand the task. Then, literature review of hand-based object manipulation interaction techniques in virtual reality was conducted to create a taxonomy of design factors. Potential design factors for hand-based interaction techniques were selected, based on which two interaction techniques were designed and further refined using pilot tests. A controlled experiment with 12 participants was conducted to evaluate these two interaction techniques of (1) push and poke and (2) custom axis with C/D gain, for the plane alignment task using pinch-based direct manipulation as a baseline. From this study, it was found that push and poke was subjectively ranked more precise and preferred because it was faster, easy to learn and easy to use and participants were confident using it. Based on the results of the study, design implications for future hand-based interaction techniques for precise plane alignment in virtual reality are discussed.

Keywords: Human-Computer Interaction, Virtual Reality, Interaction Techniques, Jaw Osteotomy Operation Planning, Object Manipulation, Hand Interaction, Experimental Research, Gesture Recognition, Fitt's Law.

Acknowledgements

I would like to take this opportunity to acknowledge everyone who contributed to completing my thesis work.

Most of this thesis work was carried out as part of my Research Assistant tenure in the Multi-Model Interaction Group (MMIG) at TAUCHI Research centre, Tampere University. I would like to thank Prof. Roope Raisamo, my supervisor for giving me this opportunity to work in this project, providing me the freedom to explore and guiding me in the right direction when needed that helped me to complete the thesis.

I would also like to thank Dr. Jari Kangas, for helping me by piloting the design and having regular discussions. All your comments and feedback have helped me to improve this work immensely.

I am grateful to Helena Mehtonen from TAYS, for having insightful, long discussions which helped me to gather knowledge and understand the context better. This helped me immensely improve this work.

I would also like to thank my wife, Devangini Patel who also helped me by piloting the design and gave critical comments that helped me to improve the design.

Finally, I would like to acknowledge the funding agency, Business Finland, project number : BF/DPI, 31220199303, due to which I was able to work on this project and produce this work.

Contents

List of Figures

List of Tables

1 Introduction

In the medical domain, medical professionals are currently diagnosing and planning operations using a traditional two dimensional (2D) monitor with keyboard and mouse interface. Using this Graphical User Interface interaction style, medical professionals perceive 3D medical data on 2D interface. This requires medical professionals to mentally reconstruct the 3D anatomical structure in 2D which can be complicated, error-prone, and requires training. Due to which, medical operation planning in this 2D interface is time-consuming and requires a high cognitive load due to a lack of 3D perception. Viewing 3D data in a 3D environment rendered by virtual reality head-mounted display reduces the 3D to 2D information loss and provides an advantage over 2D screens when it comes to perceiving and understanding 3D human anatomy (Boléo-Tomé, 1998; Steuer, 1992; Sutherland, 1968).

3D perception could be especially important for professionals working in Jaw Osteotomy Operation planning. Osteotomy is a surgical incision performed on bones to shorten, lengthen, or change their position and orientation (Di Matteo et al., 2013). Jaw osteotomies are carried for roughly 5% of the world population for jaw misalignment (a receding chin, open bite), TMJ (temporomandibular joint) disorder, sleep apnea, malocclusion problems (Posnick, 2013). In Jaw osteotomy operation planning, the cutting step consists of the following three sub-steps: (1) marking points and creating an osteotomy plane, (2) manipulating the position, orientation, and scale of the osteotomy plane (3) performing an osteotomy cut using the osteotomy plane. In jaw osteotomy operation planning, the accuracy is important since the health risk is very high at around 10%-20% (Boléo-Tomé, 1998; Shigeishi et al., 2015). In the process of jaw osteotomy operation planning, maximum accuracy is required for adjusting the osteotomy plane because this is the final step, and it compensates for the errors in the marking step. This plane adjustment step is synonymous to object manipulation interaction technique in virtual reality.

In Virtual Reality based jaw osteotomy operation planning, different interaction techniques such as tracker-based (Xia et al., 2000), 3D pen-based (Hsieh et al., 2002) and haptic pen (Olsson et al., 2015) have been used for object manipulation. Out of hand-based and controller-based methods for interaction, users prefer controller because it is more accurate and reliable (Caggianese et al., 2018; Galais et al., 2019; Gusai et al., 2017), but they would like to use hands as they don't need additional hardware (Figueiredo et al., 2018). Hand tracking is generally done through imaging-based sensors such as cameras, leap motion (Leap Motion, 2012) which is currently noisy due to the egomotion of the head, lack of FOV, occlusion, illumination, background noise (Oculus Quest, 2020). Due to these factors, hand interaction is not reliable and accurate enough to be used for jaw osteotomy operation planning. Previous research work has designed interaction techniques for manipulating 3D objects however these may or may not be applicable for manipulating a 2D plane. Thus, there is a need for designing and evaluating hand-based interaction techniques for precise plane manipulation in virtual reality.

This thesis work focuses on designing and evaluating precise and efficient hand-based interaction techniques for plane alignment in virtual reality. In the design process, the contextual inquiry was used to understand how the medical professionals adjusted the plane while performing the osteotomy operation planning. Based on this contextual inquiry, a plane adjustment task was created. Next, a literature review of the existing object manipulation interaction techniques in virtual reality was conducted to understand the different design factors. After this, potential design factors for interaction techniques for the plane alignment task were selected. Initial interaction techniques of (1) push and poke and (2) viewing angle with Control-Display (C/D) gain were designed and implemented. These two interaction techniques were iterated based on feedback from initial pilots with 2 HCI researchers. The second interaction technique was refined and renamed to custom axis with C/D gain. To evaluate these two interaction techniques for the precise plane alignment task, a controlled experiment with 12 participants was conducted. Pinch-based direct manipulation was used as a baseline. The research questions for the study were:

1. Which interaction technique is more accurate and preferred for object manipulation in virtual reality?

2. Do these interaction techniques support both small and large movements and which one(s) is required for precise object manipulation?

3. Should hand-based interaction techniques use gesture recognition?

The objective measures of task completion time, the accuracy of plane placement, number of interactions, and subjective measures of confidence, precision, learnability, usability, intuitiveness, naturalness were collected in the study.

The results of the controlled study show that there is no significant difference between the interaction techniques in terms of precision of plane placement, however, the time to complete the task using custom axis with C/D gain is significantly much higher than both pinch and push and poke. Push and poke was easy to learn due to the familiarity and naturalness of these gestures. Push and poke was easy to use in comparison to custom axis with C/D gain. Participants were more confident when using push and poke in comparison to both pinch and custom axis with C/D gain due to the usability issues with the latter. In terms of ranks of the precision rating, push and poke was first, followed by custom axis with C/D gain and finally pinch. When asked to rate the techniques based on preference, participants preferred push and poke first then pinch, and finally custom axis with C/D gain. In summary, the push and poke is faster, easy to learn and use, participants are confident using it, ranks high in subjective precision and most preferred. Custom axis is the most precise after push and poke. Pinch comes second in terms of speed, naturalness, ease of learning, preference after push and poke.

The results helps to answer the proposed Research Questions. The design implications of the study are (1) interaction techniques for precise object manipulation should support smaller movements, (2) interaction techniques could support large movements in addition to small movements for efficient object manipulation (3) interaction techniques should try to avoid gesture recognition and if not then strategies to compensate for the delay in gesture recognition and noisy hand tracking should be incorporated. The study was limited in terms of the training time provided to the participants and parameters of the physics used for push and poke interaction technique. Future work could explore designing the combination of techniques suggested by participants: (1) pinch and poke and (2) pinch and rotation handle of custom axis with C/D gain as well as adding feedback so that participants can understand when the interaction has started and stopped.

In summary, the contributions of the study are: (1) taxonomy of design factors for hand interaction techniques for object manipulation in virtual reality, (2) proposed designs for interaction techniques for plane adjustment task for performing jaw osteotomy operation planning in virtual reality, (3) empirical validation of interaction techniques for plane alignment task, (4) design implications for future hand-based interaction techniques for precise plane alignment in virtual reality.

This thesis work has nine chapters. In the second chapter, virtual reality as an interaction style is introduced and the interaction loop and hardware devices to create immersion in virtual reality are explained. In the third chapter, the anatomy of hand and process of hand tracking are discussed which helps to understand the limitations of hand tracking and hand-based interaction in virtual reality. This chapter also presents the taxonomy of design factors for hand-based interaction techniques in virtual reality. The fourth chapter introduces the context of jaw osteotomy operation planning, explains the process and discusses the existing research work in virtual reality applications for jaw osteotomy operation planning. The fifth chapter explains the design process which was carried out to create two interaction techniques of (1)push and poke and (2)custom axis with C/D gain. The sixth chapter discusses the study that was carried out to evaluate the interaction techniques for plane alignment task and to answer the three research questions. The seventh chapter presents the quantitative results from the study which are explained by supporting participants' quotes. The eight chapter discusses the findings and design implications from the study, the limitations of the study and future work. The final chapter concludes the thesis by summarizing the contributions of the work and explains the implications of this research work in a broader sense of designing hand-based interaction for virtual reality.

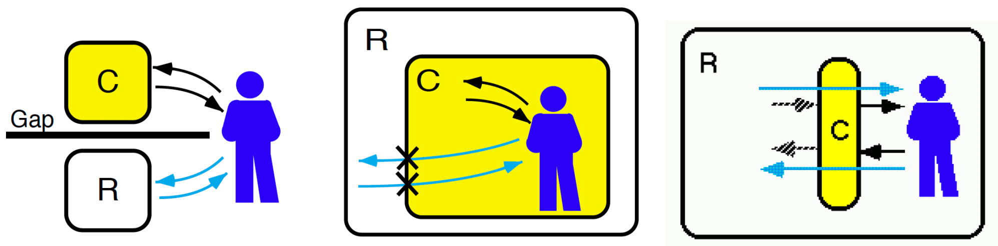
2 Virtual Reality and Related Interaction Methods

In this chapter, interaction style, and the technology behind virtual reality are explained. First, the concept of virtual reality is introduced and the interaction style for creating immersion is explained. The concept is further elaborated through a discussion of important factors in achieving objective and subjective immersion in virtual reality. Finally, the input and output devices for creating objective and subjective immersion in virtual reality are listed.

2.1 Virtual Reality

Ellis (1994) defined Virtual Reality as an "interactive, virtual image displays enhanced by special processing and by non-visual display modalities, such as auditory and haptic, to convince users that they are immersed in a synthetic space".

2D screen of a computer and mobile phone use GUI for the user to interact with the system. According to Rekimoto and Nagao (1995) user interfaces can be classified into four human-computer and human-world interaction styles. These four human-computer interaction styles include (1) graphical user interface, (2) virtual reality, (3) ubiquitous computing, and (4) augmented reality as shown in Fig. 2.1. In a graphical user interface as shown in Fig. 2.1(a), the user either interacts with the computer or the real world but not both at the same time. There exists a gap between the computer and the real world. In comparison to this, in virtual reality, as shown in Fig. 2.1(b), the computer world replaces the real world and the user interacts with the computer world only. In augmented reality, as shown in Fig. 2.1(c), information is augmented in the real world with the help of the computer and the user can interact with the real world through the computer.



(a) Graphical User Interface (b) Virtual Reality (c) Augmented Reality

Figure 2.1 *Three of the four human computer interaction styles (Rekimoto & Nagao, 1995).*

According to Milgram et al. (1995), virtual reality and augmented reality are related and can be viewed on the opposite ends of the reality-virtuality continuum as shown in Fig. 2.2. In augmented reality, the user is bound to real-world properties such as physics, time, gravity, material properties. On the other hand, immersion of user in a simulated world where the real-world properties may not be followed is virtual reality. Virtual reality is beneficial for reducing the cost of setting up physical space to recreate a scenario of the real world or creating a new scenario that is not possible in the real world.

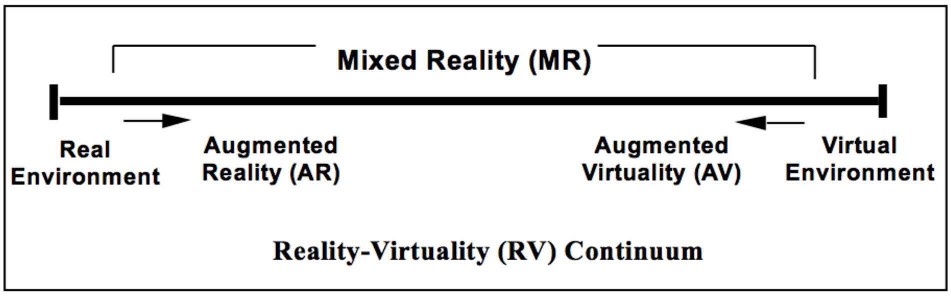


Figure 2.2 *The reality-virtuality continuum proposed by Milgram et al. (1995).*

2.1.1 Immersion and Presence in Virtual Reality

Lingard (1995) defined three stages of virtual reality systems: the passive, exploratory, and immersive stages. In passive systems, the user is only viewing the virtual reality content. In exploratory systems, the user can navigate around the virtual reality world. In immersive systems, the user can sense the virtual world through multiple modalities and interact with the virtual objects. One believing that he/she is physically present while perceiving simulated world is immersion. Immersion in virtual reality is important for users to believe that the virtual world is physically real. The immersion of a user in the virtual world can be measured through immersion and presence. Immersion is the objective fidelity of the virtual reality system whereas presence is the user's subjective response to the immersion of the virtual reality system (Slater, 2003). According to Slater (2003), immersion consists of several factors such as sensory fidelity including fields of view, resolution, behavioral fidelity of human ability in the virtual world, system latency, and other physical properties from the real world such as temperature, airflow, gravity. Bowman and McMahan (2007) broke down visual immersion into further factors such as field of view (FOV), size of the display, the field of regard (FOR), resolution of the display, stereoscopic vision to provide depth cues, head tracking, frame rate, realistic light and refresh latency.

Presence is a complex term and the factors contributing to presence are not agreed upon yet. Some of these factors are involvement, control, the anticipation of events, awareness, the meaningfulness of experience which are grouped into four parts of distraction, control, sensory, and realism factors (Witmer & Singer, 1998).

According to Steuer (1992), the important factors that contribute to the presence in virtual reality are (1) vividness or realness and (2) interactivity as shown in Fig. 2.3. Vividness or realness refers to the different modalities in the experience and the resolution of these different modalities. Interactivity is the extent of the user to interact with the virtual world and objects and change their form. Vividness is divided into breadth and depth. Breadth refers to the different modalities supported by different devices in the experience and depth refers to the resolution of these devices. Interactivity is further divided into speed, range, and mapping. Speed refers to the speed at which interactions can be processed by the system. The range refers to the different interactions supported by the system. Mapping refers to the process in which the input is converted to an interaction in the virtual world. (Steuer, 1992) According to Slater (2003), the realness and interactivity of virtual reality systems should be designed considering the perception and motor systems of users as human physiology can not be changed.

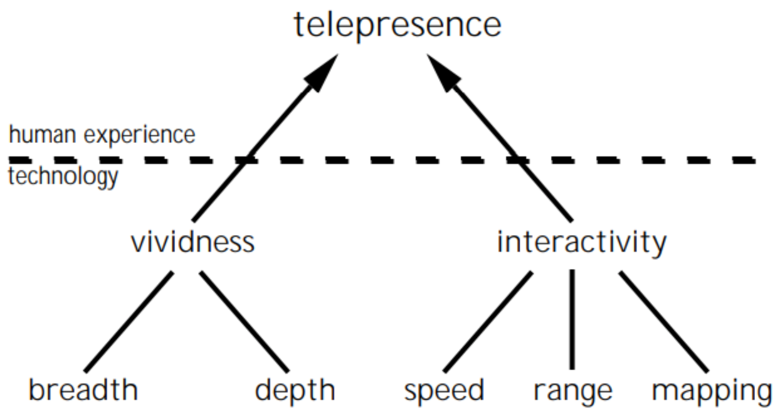


Figure 2.3 *The dimensions of immersion (Steuer, 1992).*

2.2 Virtual Reality Technology

Norman (2013) has proposed the action cycle to explain how the user interacts with a system. According to the action cycle, the user goes through 7 stages as shown in Fig. 2.4. Users first evaluate the world by perceiving, interpreting, and evaluating the desired action to take. After evaluating, the user will try and execute that action by performing the sequence of action that was intended. In the case of virtual reality system, the user interacts in the 3D world simulated by the computer system through the interaction loop as shown in Fig. 2.5.Input and output devices are used for creating realness and interactivity for virtual reality applications.

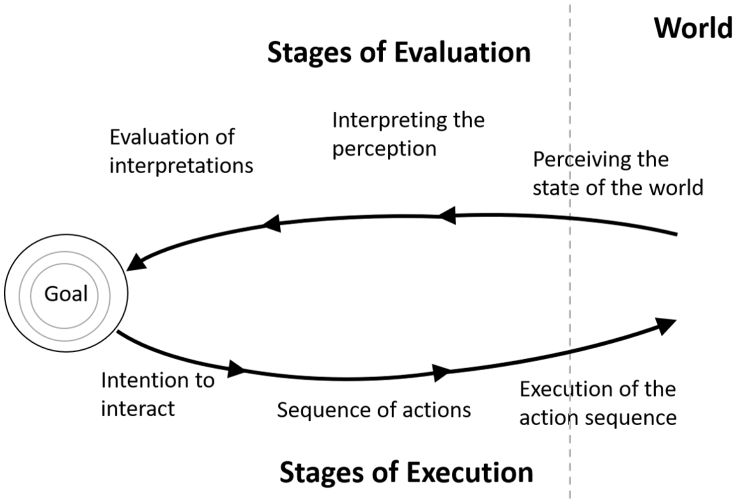


Figure 2.4 *The 7 stages of Don Norman's Action cycle. (Hermann & Weber, 2009).*

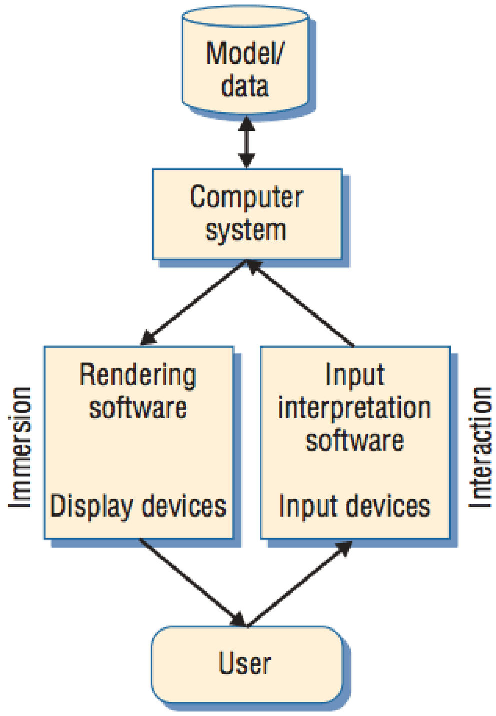


Figure 2.5 *The interaction loop in virtual reality systems (Bowman & McMahan, 2007).*

Sutherland (1965) described the concept of ultimate display in which a user can be in a simulated world which it feels like a real world. Immersion in the virtual world can be experienced by a user through the head-mounted display through the five basic human senses such as realistic visual, 3D audio, haptic feedback, smell, and taste (Sutherland, 1965). Natural interaction with objects in virtual reality also contribute to this immersion (Sutherland, 1965). The current technology is not yet completely capable to provide the experience of immersion as described in the concept of ultimate display. Haptics is one of the areas which needs to be developed further for providing real-world physical properties in virtual reality, such as sense of weight, touch, force feedback, to achieve the vision of ultimate display.

Several hardware devices are needed to support realness and interactivity in virtual reality. The following sections describe the input and output devices that are used for creating realness and interactivity in virtual reality.

2.2.1 Virtual Reality Output

To achieve immersion in virtual reality, the devices should facilitate the input modalities of humans. Humans can see the world, hear, feel through touch, smell, and taste. Some of the output devices that are commonly used in virtual reality are Head Mounted Display (HMD), audio output with spatial localization, and vibrotactile feedback from virtual reality controllers.

Head Mounted Displays (Sutherland, 1968), CAVE (Cruz-Neira et al., 1992), chameleon virtual reality (Fitzmaurice et al., 1993), fish tank virtual reality (Ware et al., 1993) are some of the approaches to display virtual reality content. CAVE (Cruz-Neira et al., 1992) is a wall that has content projected on it. The user's position is tracked and the content on the wall is rendered relative to the user's position. Chameleon virtual reality (Fitzmaurice et al., 1993) is a handheld visual device that can be translated in space to perceive the virtual reality content. Fish tank virtual reality (Ware et al., 1993) is when the display is stationary and the content changes in relation to the tracked user position, orientation, and viewing angle. The head-mounted display (HMD) is the most used approach currently. As the name suggests, HMDs are visual displays that are strapped onto the human eyes and the user position, orientation is tracked in the space. Based on the user's position and orientation in the virtual world, content on the display would change relatively. In (Pausch et al., 1993), it is shown that head-tracked virtual reality display created a much better internal representation of the virtual space. Varjo VR-3[1], HTC vive[2], Oculus Quest 2[3], Valve Index[4] are some of the commercially available modern head mounted displays. These displays used in tracking the user's head position and rotation in the virtual space. Inside-out and outside-in are the two methods that are used in tracking the user position. Inside-out tracking uses a combination of accelerometer, gyroscope, multiple calibrated cameras mounted on the exterior of the head-mounted display and runs visual-SLAM to calculate the position and orientation of the display. In the outside-in tracking method, there are external sensors such as laser sweepers or time of flight sensors that track marker

1 `https://varjo.com/products/vr-3/`
2 `https://www.vive.com/eu/product/`
3 `https://www.oculus.com/quest-2/`
4 `https://www.valvesoftware.com/en/index`

patterns embedded on the head-mounted display exterior. Using triangulation of these tracked points the position and orientation are estimated.

Audio or sound in virtual reality adds a new dimension to experience immersion in virtual reality. The head-related transfer function (HRTF) is used to generate 3D spatial audio, where the user can perceive the spatial audio using their headphones to hear the sound as if it is played from a specific point in a 3D space[5]. This output system can be used to create spatial audio localization cues to grasp human attention. This can also be used as a feedback modality that can work along with visual interaction to create immersion.

One of the most comment approach for providing haptic feedback is vibrotactile feedback from controllers. Phantom series of 3D systems Inc.[6], the Omega and Delta series of Force Dimension[7] are some of the commercially available desktop force feedback devices that are available. These devices can simulate physic properties such as rigid body, elasticity, gravity confined to a small, fixed space. CyberGrasp[8], H-glove[9], Dexmo[10], Haptx[11], Plexus[12], and vrgluv[13] are some of the commercially available force-feedback haptic gloves. These gloves can be used in feeling the shape, size, texture, stiffness of virtual objects. Some of these gloves have both force and tactile feedback. The limitation of these haptic glove devices is that it is not suitable for all scenarios such as pulling a lever or lifting a weight. Tesla suit[14] is a full-body haptic feedback suit that provides tactile and thermal feedback throughout the body. Ultraleap[15] is an ultrasonic-based device that enables us to feel the sense of touch on hands in mid-air (Sand et al., 2015).

Smell and taste are the least explored senses in virtual reality. There are a few devices to create digital smell and taste for the users in virtual reality. The smell could be produced using mechanical systems to diffuse molecules in the air (Dmitrenko et al., 2017). Taste could be created by electrical stimulation of the tongue (Nakamura & Miyashita, 2011) or thermal simulation of the mouth (Cruz & Green, 2000) and nose (Suzuki et al., 2014).

5 `https://developer.oculus.com/learn/audio-intro-spatialization/`
6 `https://www.3dsystems.com/haptics/`
7 `https://www.forcedimension.com/`
8 `http://www.cyberglovesystems.com/cybergrasp`
9 `https://www.haption.com/fr/products-fr/hglove-fr.html`
10 `https://www.dextarobotics.com/`
11 `https://haptx.com/`
12 `http://plexus.im/`
13 `https://www.vrgluv.com/enterprise`
14 `https://teslasuit.io/`
15 `https://www.ultraleap.com/`

2.2.2 Virtual Reality Input

Input devices serve as means of interaction in Virtual reality. The most basic input system used in virtual reality is an inertial sensor that is used by head-mounted display or controllers to estimate the 3D orientation and position and synchronize with the computer graphics that is rendered. These inertial sensors are not accurate and so in most cases, they are combined with other tracking techniques such as inside-out tracking or lighthouse, or outside-in tracking. (Strickland, 2007)

Virtual reality controllers are the most used input device. HTC vive[16], oculus quest controller[17] are some of the commercially available controllers. These controllers are tracked in 6 DoF and have various buttons such as trigger and joystick. Valve index controllers[18] has additional sensors to track the finger position while the controller is strapped to the hand.

2D mice are best suited for 2D displays and not suited for Virtual reality as there only 2 Degree of Freedom (DoF) available (Kim & Choi, 2019; Santos et al., 2009). The 2D mouse can be translation in X and Y axis to rotate virtual mouse pointer in X and Y in virtual reality. This input can also be coupled with three buttons and a scroll option available in the mouse. 3D desk mouse[19] are also available that can be used as an input in VR.

3D virtual reality pens are very similar to controllers where they are tracked in 3D space and in addition, there is a pressure sensor on the tip of the pen which can be used to interact with real-world surfaces where the pen can be pushed like using a normal pen. Logitech VR ink[20], Wacom VR pen[21], VR free 3D stylus[22], Massless pen[23] are some of the commercially available 3D Virtual reality pens. There are other buttons and touchpads that are also available for input. (Pham & Stuerzlinger, 2019) says users prefer using a 3D pen over a controller for object selection in virtual reality.

3D probes are mechanical arms with 6 DoF joints and the position and orientation of the tip are calculated using kinematics. The user can hold the end effector and operate the system which also has buttons. Custom end effectors can also be attached to these systems. Phantom series of SensAble Inc.[24], Polhem Haptic De-

16 `https://www.vive.com/us/accessory/controller/`
17 `https://www.oculus.com/quest-2/`
18 `https://www.valvesoftware.com/en/index/controllers`
19 `https://3dconnexion.com/uk/spacemouse/`
20 `https://www.logitech.com/en-gb/promo/vr-ink.html`
21 `https://developer.wacom.com/en-us/wacomvrpen`
22 `https://www.sensoryx.com/products/vrfree-3d-stylus/`
23 `https://massless.io/`
24 `https://www.3dsystems.com/haptics/`

vice of Forsslund Systems AB[25], the Omega and Delta series of Force Dimension[26] are some of the commercially available 3D probes that can be used as an input in Virtual Reality.

Gaze is one of the input systems that allows the system to know what the user is seeing or interested in the virtual scene. Tobii VR[27], Pico Neo 2 Eye[28], Varjo VR3 pro[29], Vive pro eye[30] are some of the commercially available VR devices that come with eye-tracking. (Pfeuffer et al., 2017) shows that adding gaze to the hand tracking pinch increase the accuracy in object selection which in turn increases productivity.

Speech can also be used as an input in Virtual reality. Studies show that speech is 3 times faster than using a keyboard for text entry (Ruan et al., 2018). Unity speech recognition API, Microsoft windows speech recognition are some of the commerically available speech recognition software that can be used. The currently available solutions have limitations such as inconsistency in recognition, works only with low background noise and the dialog and commands should be modeled manually and it is limited to the list of commands (Blackley et al., 2019).

Blowing air can also be used as an input in Virtual reality where the strength of airflow can be translated into different actions. (Cruz Cebrian, 2017; Sra et al., 2018) are some of the research that has been carried out to study blowing air as an input.

Quite recently brain-computer interface (BCI) devices have emerged in the market. HTC Vive, OpenBCI, and Tobii are working together to develop a BCI based Virtual Reality HMD called Galea[31],[32]. NextMind[33] is one of the commercially available BCI interfaces through which user can interact in Virtual Reality. Using BCI for interacting in Virtual reality is a non-invasive method, but the technology is still far from reality and currently, only limited actions can be recognized.

Full-body tracking can be achieved using Vive tracker[34] or Tundra Tracker[35] worn on the body which is tracked. The human avatar can be mapped to these trackers. Tesla suit[36], Kinect[37] and optical motion capture methods (Cao et al., 2019; Xu et al., 2019) can also be alternative techniques to achieve full-body tracking. Vive

25 https://www.forsslundsystems.com/
26 https://www.forcedimension.com/
27 https://vr.tobii.com/
28 https://www.pico-interactive.com/neo2.html
29 https://varjo.com/products/vr-3/
30 https://www.vive.com/eu/product/vive-pro-eye/overview/
31 https://galea.co/
32 https://www.tobii.com/group/news-media/press-releases/2021/2/tobii-valve-and-openbci-engaging-in-research-collaboration-to-make-vr-gaming-more-immersive/
33 https://www.next-mind.com/
34 https://www.vive.com/eu/accessory/vive-tracker/
35 https://www.tundratracker.com/
36 https://teslasuit.io/
37 https://azure.microsoft.com/en-us/services/kinect-dk/

facial tracker[38] can also be used as an input system to track facial expressions of the user and used for social applications.

Recently, users using their own hand as an input in Virtual reality has become more common. Leap motion (Leap Motion, 2012), Kinect[39], Oculus quest hand tracking (Oculus Quest, 2020), Vive hand tracking[40], Varjo[41] are some of the virtual reality devices that can recognise and track human hand using optical camera. Sec. 3 discussed more in detail about the hand based interaction techniques. Hand recognition based interaction is not reliable and accurate (Caggianese et al., 2018; Galais et al., 2019; Gusai et al., 2017) due to occlusion, the egomotion of the head, lack of FOV of the hand tracking sensor (Mine et al., 1997). Hand can also be tracked using a wearable glove such as CyberGrasp[42], H-glove[43], Dexmo[44], Haptx[45], Plexus[46], and vrgluv[47]. These wearable gloves can track the hand more accurately compared to hand recognition-based solution but it requires additional hardware to be purchased.

38 `https://www.vive.com/eu/accessory/facial-tracker/`
39 `https://azure.microsoft.com/en-us/services/kinect-dk/`
40 `https://developer.vive.com/resources/vive-sense/sdk/vive-hand-tracking-sdk/`
41 `https://varjo.com/products/vr-3/`
42 `http://www.cyberglovesystems.com/cybergrasp`
43 `https://www.haption.com/fr/products-fr/hglove-fr.html`
44 `https://www.dextarobotics.com/`
45 `https://haptx.com/`
46 `http://plexus.im/`
47 `https://www.vrgluv.com/enterprise`

3 Hand based Interaction Techniques for Object Manipulation in Virtual Reality

In this chapter, the technology of hand tracking and a summary of hand-based interaction techniques for object manipulation are discussed. First, the anatomical structure of the human hand is introduced as it helps in understanding how the human hand can create poses. Then, the different methods of hand tracking are discussed as it helps to understand the limitations of hand tracking that the interaction techniques must deal with. The different interactions in virtual reality are introduced and the interaction of object manipulation is discussed in detail. Finally, a taxonomy of design factors for hand-based interaction techniques for object manipulation in virtual reality is presented.

3.1 Anatomy of Hands

The functional capabilities of the human hand depend on the anatomical structure of the human hand (Schwarz & Taylor, 1955). The human hand consists of 27 bones and 27 joints as shown in Fig. 3.1. The wrist contains 8 carpal bones; the palm has of 5 metacarpal bones. The fingers consist bones called phalanges; the thumb contains 2 phalanges, and the other fingers have 3 phalanges each. The wrist bones connect with the radius and ulna bones to form the wrist joint. The phalanges join with the carpals to form the metacarpophalangeal joints. The phalanges join each other at interphalangeal joints. These metacarpophalangeal and interphalangeal joints work like hinges. The human hand has 34 muscles. The bones along with the muscles and stability of ligaments allow the hand to form various hand poses to perform various tasks (Panchal-Kildare & Malone, 2013).

3.2 Hand Tracking Techniques

Hand tracking is a technique to determine the 3D pose of the hand. A sensor is generally used for hand tracking. The most common sensors used for hand tracking include RGB camera, depth camera, Time of Flight (ToF) camera, infrared camera, stereo vision cameras, wearable gloves, etc.

For sensors, other than wearable gloves, computer vision algorithms are used to perform hand tracking in real-time. The first step in the computer vision algorithm is to detect the region of interest (ROI) in the image in which the hand is present (Sharp et al., 2015). The image is later cropped based on the ROI so that a minimal part of the image is processed. Then, the hand pose is estimated from the cropped image.

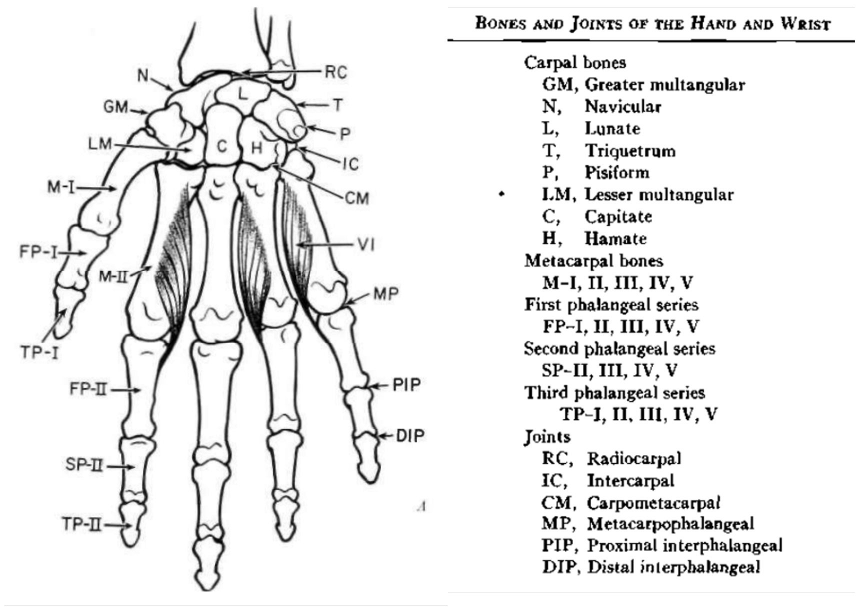
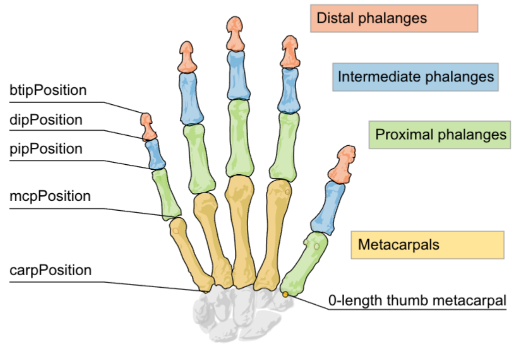


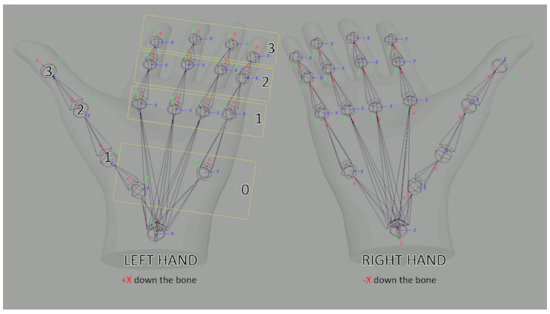
Figure 3.1 *The anatomical structure of the hand including bones and joints. Image taken from (Schwarz & Taylor, 1955)*

There are two types of traditional computer vision-based hand pose estimation techniques: (1) appearance-based and (2) 3D model-based. In appearance based approach, the hand pose is predicted based on the visual features such as intensity values (Lanitis et al., 1995), contours (Cootes et al., 1995; Lanitis et al., 1995), histograms (Freeman & Roth, 1995), moments (Schlenzig et al., 1994), and fingertips (Ahmad & Tresp, 1993). In this approach, a limited set of hand poses are used for training and hence this approach can predict a discrete set of hand poses. However, due to this limitation, this technique is rather fast. In the 3D model-based approach, an initial hypothesis of the hand pose of the 3D model is created which is later refined iteratively using optimizations techniques such as Iterative Closest Point (ICP) and particle swarm optimization (PSO) to reduce the cost function (Oikonomidis et al., 2011; Sharp et al., 2015). The 3D model is based on the anatomical model of the hand. This technique produces a continuous range of hand poses; however, this technique is computationally expensive.

Recently, deep learning techniques for predict the hand pose have been created. They techniques use either one or several networks including encoders (Boukhayma et al., 2019), CNN (Boukhayma et al., 2019; Ge et al., 2019; Oberweger & Lepetit, 2017), residual network (Ge et al., 2019; Oberweger & Lepetit, 2017; Wan et al., 2018), segmentation networks (Zimmermann & Brox, 2017), 2D to 3D projection (Wan et al., 2018; Zimmermann & Brox, 2017). Oculus Quest and Leap motion use deep learning techniques which predict 3D points of joints of a 3D hand model (Leap Motion, 2012; Oculus Quest, 2020). Their hand models differ slightly. The hand model used by Oculus Quest is shown in Fig. 3.2(b) and the hand model used by Leap Motion is shown in Fig. 3.2(a).



(a) The 3D hand model used in Leap Motion. Image taken from (Leap Motion, 2012)



(b) The 3D hand model used in Oculus Quest. Image taken from (Oculus Quest, 2020)

***Figure* 3.2** *3D hand models used by Leap Motion and Oculus Quest.*

Hand tracking is challenging due to the DoF of hands, and the variations in hand size and shape (Sharp et al., 2015), occlusion, illumination, background noise (Oculus Quest, 2020).

3.3 Interactions in Virtual Reality

Hand tracking is used in virtual reality so that users can interact in the virtual world. According to Bowman and Hodges (1999), interaction in virtual reality can be broken down into three main types: (1) wayfinding, 2) navigation, and (3) object selection and manipulation. In wayfinding, the user can locate themselves in the virtual world. In navigation, the user can move from one location to another. In object selection and manipulation, the user can select a target object and transform the object by changing the position, orientation, or scale of the object. In this thesis work, object manipulation interaction is of interest because medical professionals will be manipulating the skull and/or osteotomy plane for jaw osteotomy operation planning.

3.3.1 Object Manipulation

Object manipulation is the process of changing the translation and rotation of an object, optionally the scale and shape of the object in addition (Bowman & Hodges, 1999). Bowman and Hodges (1999) studied the task taxonomy of object selection and manipulation as shown in Fig. 3.3. Object manipulation requires that the object is selected first, then manipulated, finally released. Object selection tasks consist of feedback, an indication of an object, and an indication of the select operation. Object manipulation task consists of attaching the object to hand or gaze, changing the position or orientation, and providing feedback. The object release task consists of an indication of drop and feedback of the operation.

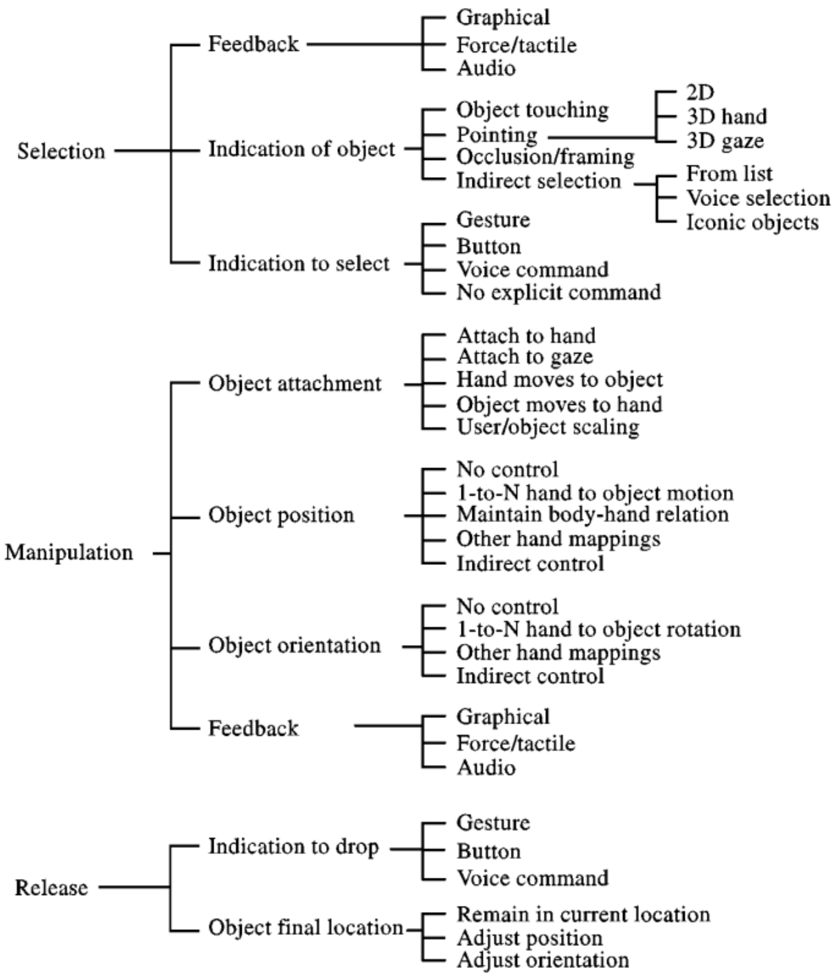


Figure 3.3 *The task taxonomy of object selection and manipulation. Image taken from (Bowman & Hodges, 1999)*

3.3.2 Fitts's Law

Fitts's Law (Fitts, 1954) is a principle relevant for object selection and it helps to determine the measure for difficulty of object selection task. Fitts's law states that the total movement time is a measure of difficulty which is a logarithmic ratio between the distance to the target object and object size as shown in Eqn. 3.1. This means if a user wants to reach an object placed at a distance, the difficultly to select the object is greater when the object size is smaller and difficultly is lesser for selecting a bigger object. This was initially proposed in one dimension. Later in (Accot & Zhai, 2003), the same Fitts's law for target pointing task was explored in 2 dimensions and later on many other studies examined Fitts's law in a three-dimensional environment and found the law holds true (Mateo et al., 2005; Murata

& Iwase, 2001). Fitts's law also holds true in object manipulation in virtual reality (Y. Wang & MacKenzie, 1999). (Graham & MacKenzie, 1996) study shows that hand movements for object selection can be separated into two phases: (1) initial fast and imprecise movement and (2) final slow and precise movements.

$$ID = log_2\left(\frac{2D}{W}\right) \tag{3.1}$$

where:

$$\begin{aligned} ID &: \text{index of difficulty,} \\ D &: \text{distance to the target} \\ W &: \text{the target width} \end{aligned}$$

3.4 Taxonomy of design factors for hand interaction techniques for object manipulation in virtual reality

Several interaction techniques (Bossavit et al., 2014; Cho & Wartell, 2015; Kruger et al., 2005; Mendes et al., 2017; Mlyniec et al., 2011; Poupyrev et al., 1996; Song et al., 2012) have been developed for the controller and hand-based object manipulation in virtual reality. These techniques usually differ in terms of design factors. Fig. 3.4 shows the taxonomy of design factors for the controller and hand-based interaction techniques for object manipulation in virtual reality. Due to limited work on hand-based interaction techniques for object manipulation in virtual reality, controller-based interaction techniques were also included in the literature review as controller-based interaction techniques could be adapted to hand-based.

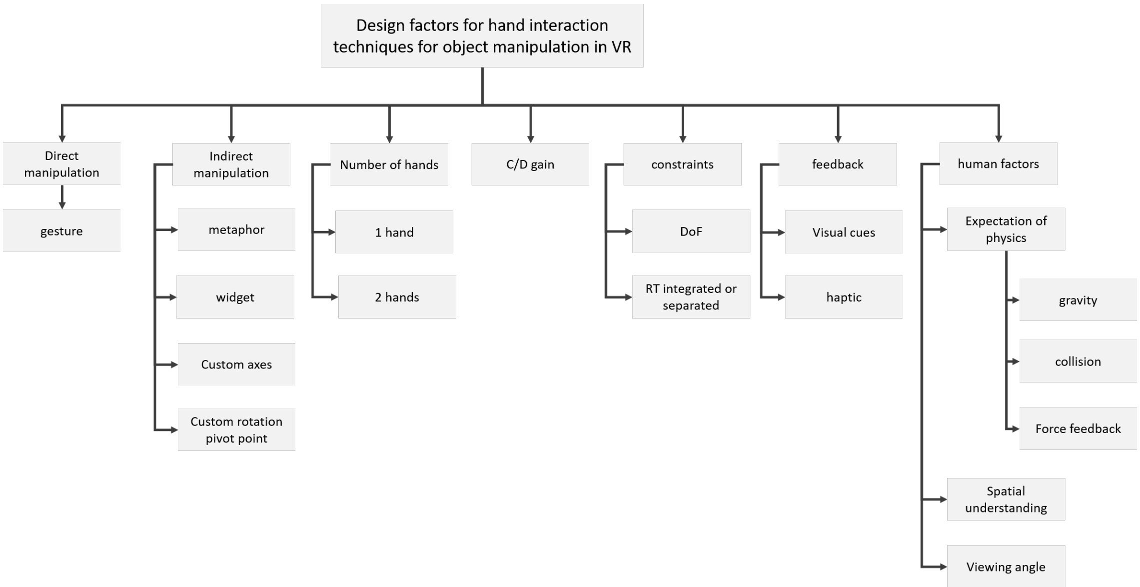
These design factors along with the related work using these design factors have been discussed in detail below.

3.4.1 Direct manipulation

Direct manipulation coined by Shneiderman (1981) is an interaction technique in which objects are interacted physically, incrementally, reversibly, with immediate feedback. Human hands are the input device for direct manipulation in VR (Jacoby et al., 1994).

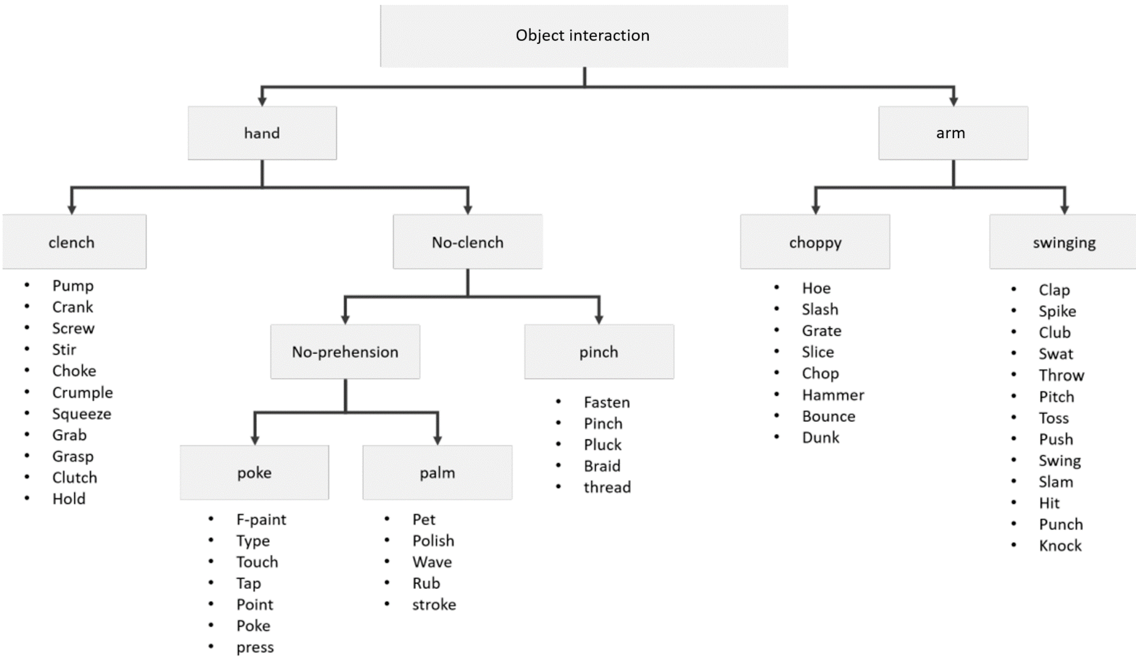
In virtual reality, generally, pinch and grasp gestures are used for direct manipulation (Caggianese et al., 2018; Galais et al., 2019; Gusai et al., 2017). However, several other gestures also exist. Klatzky et al. (1993) reviewed and created a taxonomy of different interaction gestures with objects in the real world. These hand and arm gestures are shown in Fig. 3.5. Thus, there are several hand gestures that

Figure 3.4 *Taxonomy of design factors for interaction techniques for object manipulation in virtual reality*



can be explored for object manipulation in VR.

Figure 3.5 *Taxonomy of gestures for object interaction, Image adapted from (Klatzky et al., 1993).*



3.4.2 Indirect manipulation

In indirect manipulation, the hand movements are mapped and transformed into the operations in the VR space using metaphors and widgets. These metaphors and widgets provide affordances and signifiers to the user to make them easier to use.

Metaphors

Metaphor is a way of using analogy to help users create a mental model of the interaction technique (Erickson, 1995). Go-Go Interaction Technique (Poupyrev et al., 1996) uses the metaphor of growing arms to overcome the limitation of physical reach of hands.

Handlebar technique (Song et al., 2012) has been used for manipulating single and multiple objects using two hands. The handlebar's position changes immediately based on the positions of the two hands while the hands are performing pointing gestures. The object manipulation mode is activated when the hands are closed to a tight grip and the object(s) along the handle are selected. The user can translate and rotate the object by moving the handlebar with two hands. The user can uniformly scale the object by changing the distance between the two hands. The evaluation of this technique showed that the handlebar provided a strong sense of control to the user, the handlebar metaphor provided an intuitive way to learn, however, users experienced fatigue with this technique.

MAiOR (Mid-Air Objects on Rails) (Mendes et al., 2017) is a controller-based interaction technique that offers both 3-DOF and 1-DOF manipulations for translation and rotation. The user must press a button on the controller to create custom axes for the operation. The custom axes act like a rail and the object is constrained to translate or rotate along that axis. In their evaluation against widgets and direct manipulation, this technique was more accurate than direct manipulation but less than widgets, however, this metaphor was hard to remember.

The spindle technique (Mlyniec et al., 2011) creates a line between the two hands and the centre of the line representing the centre for rotation and scaling. The spindle improved the understanding of the interaction. Their evaluation showed that this technique was faster than the one-handed wand technique and mouse interaction.

Spindle Wheel technique (Cho & Wartell, 2015) was created using button ball devices. The spindle is created when the button balls are activated. The translation and scaling operations work similarly to the handlebar. The wheel is created when one of the hands starts rotating like a wheel. This rotation movement rotates the object in terms of either yaw or roll. This technique was compared with Spindle, a one-handed version and a version with scaling operated by one ball and direct manipulation by the other ball. They found that this method was faster and more preferred than Spindle. One-handed and scaled versions were faster than the original version and users preferred using the one-handed versions.

In the Crank Handle technique (Bossavit et al., 2014), a crank handle is created along an axis of the object which is closest to the dominant hand. C/D gain is applied to scale the movement. The gain factor depends on the speed of the rotation. They compared it with touch-screen-based Grasping Object (Kruger et al., 2005) and Handlebar technique (Song et al., 2012), they found that this technique performed similar to Handlebar in terms of accuracy, time taken and precision.

Paper metaphor-based technique (R. Wang et al., 2011) allowed the users to mimic paper rotation along fixed x, y, z axes. They compared this approach against mouse and found that this technique could save users time however it was not precise.

Widgets

Smart Pin is a widget designed for one hand interaction. The user can activate one of the operations: rotation, and scaling by grabbing one of the caps or translation by grabbing centre of the object. On grabbing, these caps are expanded to show the activation of the mode. They compared this technique with the handlebar technique and observed no difference in terms of task completion time and learnability. However, smart pin had higher ease of use, more hands' coordination, higher preference and more physical comfort as it reduces the need for large motions. (Caputo et al., 2018)

Mendes et al. (2016) created a widget for DoF separation. When compared to Precise and Rapid Interaction through Scaled Manipulation (PRISM) (Frees & Kessler, 2005) and direct manipulation, this widget helped users to do fine movements and also avoid unnecessary additional actions. The main reason was because this widget was able to independently perform translation and rotation in a given time.

Nguyen et al. (2014) designed a 7 Handle manipulation technique which has 7 points on an object model. The first three points correspond to the vertices of a triangle around the object. The next three points are the midpoints of the edges of the triangle. Changing the midpoint would adjust the neighbouring vertices of the triangle. The last point corresponds to the centre of the object. Manipulating the last point is equivalent to direct manipulation. They compared this technique with direct manipulation and found that this technique takes more time than direct manipulation. There was no significant difference in terms of intuitiveness, ease of use, preference, but this technique was better in terms of fatigue and efficiency.

Custom axes

Several widgets and metaphors (Bossavit et al., 2014; Caputo et al., 2018; Mendes et al., 2016; R. Wang et al., 2011) have fixed axes aligned with the object axes. This forces the user to perform the operations along those specific axes. There are some techniques that allow the user to create a custom axis for translation and rotation. In Handlebar technique (Song et al., 2012), the user can create a custom rotation and translation axes, by moving their hands while in pointing gesture. The evaluation of this technique showed that the handlebar gave the users a strong control. MAiOR (Mid-Air Objects on Rails) technique (Mendes et al., 2017) allows the user to create custom axes by pressing a button on the controller. In their evaluation against widgets and direct manipulation, this technique was more accurate than direct manipulation but less than widgets, however, this metaphor was hard to remember.

Custom rotation pivot point

Most interaction techniques use the centre of the object as the rotation pivot point. 7 Handle technique (Nguyen et al., 2014) allows for multiple rotation pivot by using 6 points of the triangle around the object that the user can use for controlling the rotation while the points on the opposite of the triangle remain stationary. If the first level point (vertex) is rotated, then the opposite side of the triangle is kept still. If the second level point (midpoint of edge) is rotated, then the opposite vertex of the triangle is kept still. In these ways, this approach allows the user to create multiple rotation pivot points on the object. However, in this approach, users can not create other custom pivot points.

3.4.3 Number of hands

Cutler et al. (1997) observed that users generally performed two hands interaction for object manipulation in virtual reality. However, no formal evaluation of their techniques was performed.

According to Guiard (1987), the two hands have different roles in a two handed operation. He created a framework to explain the usage of the two hands in actions: (1) right hand performs the action while the left hand acts as spacial reference to it, (2) the amount of motion differs between the two hands: the right-hand moves less distance and more times than the left hand; the left hand is used for gross movements and the right hand is used for precise movements, (3) the left-hand initiates the action.

Handlebar (Song et al., 2012) technique uses two hands. The one-handed Crank Handle (Bossavit et al., 2014) performed similarly to Handlebar in terms of accuracy, time taken, and precision. Smart Pin (Caputo et al., 2018) used one hand and it took the same time as 2 handed Handlebar (Song et al., 2012) technique. However, participants felt higher ease of use, more physical comfort, and had a higher preference for the smart pin in comparison to the handlebar. There is no clear indication of whether using one or two hands is beneficial.

3.4.4 C/D gain

C/D gain has been used for object manipulation interaction for various purposes.

Precise and Rapid Interaction through Scaled Manipulation (PRISM) technique (Frees & Kessler, 2005) adjusts C/D gain ratio for object selection and manipulation so that the movement of the VR object is less sensitive to the physical movements of the hands. Their evaluation found that PRISM provided a higher degree of precision than direct manipulation with less completion time but higher learnability.

Mendes et al. (2016) created a widget for DoF separation. They used a C/D gain factor of 0.25. PRISM (Frees & Kessler, 2005) method and direct manipulation was compare with this widget. They identified that users were able to make fine adjustments and were able to avoid unnecessary movements. They also identified user were able to perform accurate movements with scaled translations but were confused when using scaled rotations.

Osawa (2008) designed one and two-handed interaction techniques with position and viewpoint adjustments. The position moved is adjusted by a scale factor and viewpoint is also adjusted by a scale factor. These are modes that are activated when the speed is lower than a lower threshold until the speed exceeds the higher threshold. They compared direct manipulation with scaled movements and the combination, considering both one and two-hand conditions. They found that two-handed techniques with scaled movement and viewpoint are better than all other combinations however there was no usability gain as the users got confused by the automatic mode switching.

3.4.5 Constraints

Constraint added in indirect manipulation for DoF separation. Generally, these interaction techniques separate translation and rotation into two separate operations. These reduce the DoF and provide the user with more control on the object manipulation (Cutler et al., 1997). Mendes et al. (2016) created a widget for DoF separation. Compared to PRISM (Frees & Kessler, 2005) method and direct manipulation users were able to make fine adjustments and were able to avoid unnecessary movements as they were able to do translation and rotation separately. Users were able to do precise movements with scaled translations. Scaled rotation confused the users. These axes were not custom specified by the user but were aligned with the object axes.

3.4.6 Feedback

Fitts's Law states that the metrics of difficulty for a user selecting a target can be measured by a logarithmic ratio of distance from the target location and the target size(Fitts, 1954). Studies have shown that having haptic feedback can help in reducing the difficulty of this task (Corbett et al., 2016).

Various visual feedback methods for object grasping task was studied by Vosinakis and Koutsabasis (2018). It was found that users prefer to have some visual feedback than no visual feedback. The different visual feedback that they studied includes changing the colour of the object, drawing a line to the object, creating a halo effect around the object, and using shadows. They found object colouring and halo effect are the most preferred visual feedback. Drawing a line to the object was distracting (Vosinakis & Koutsabasis, 2018). (Canales & Jörg, 2020) showed that users preferred audio feedback for object manipulation more than visual feedback.

Position and viewpoint adjustments technique (Osawa, 2008) used viewpoint adjustment which magnifies the scene so that the user can precisely place the object. Silk cursor (Zhai et al., 1994) shows occlusion cues along with a transparent volume tracking point. They compared it with a wireframe cursor, and it was more accurate and faster.

3.4.7 Human factors

These human factor related to the expectation of interaction in virtual reality. Arora et al. (2019) found that all the participants were directly interacting with the object and they implicitly expected physics to be implemented within the system, specifically gravity, deformability, and contact modeling. They also found that different gestures were used based on the different contexts of use.

Viewing objects from different views helped the users to decide what action to perform, and when the viewing object is held in the user's hand the user can quickly view the object from different views (Mine et al., 1997).

Singh et al. (2021) showed that there is a cognitive conflict between the visual system and the proprioception when selecting an object in virtual reality. This conflict is created due to the hand movement velocity and hand tracking error. They also found a specific pattern when users selected the object. Users first accelerated fast to the target and then decelerated before touching the object. This corroborates with (Graham & MacKenzie, 1996) in which they showed that Fitts's law can be divided into two parts; in the first part the user moves fast to the target and in the second part, the user slows down for more precise movements to reach the target.

4 Virtual Reality Methods for Jaw Osteotomy Operation Planning

In this chapter, the context of jaw osteotomy operation planning is introduced. The process of jaw osteotomy operation planning carried out using conventional 2D-screen-based software tools is explained. This section also reviews the existing VR technologies for jaw osteotomy operation planning.

4.1 Jaw Osteotomy

Osteotomy is a surgical incision performed on bones to shorten, lengthen, or change their position and orientation (Di Matteo et al., 2013). Jaw osteotomies are performed for roughly 5% of the world population for different issues with the jaw such as jaw misalignment (a receding chin, open bite), temporomandibular joint (TMJ) disorder, tumours, sleep apnea, malocclusion problems (Posnick, 2013). If the jaw osteotomy operation is not performed, it might lead to bone degeneration and end-stage disease for some patients (Clohisy et al., 2009). The diagnosis, planning, and treatment phases of a jaw osteotomy involve several health professionals including an orthodontist, a radiologist, a surgeon, a periodontist, a prosthodontist, a dentist, and a doctor (Posnick, 2013). In the planning phase, these different health professionals take care of the various aspects of the operation such as (1) resolving jaw issues, (2) maintaining a proper airway, and (3) checking the aesthetic looks of the healed face after the operation (Posnick, 2013).

4.2 Process of jaw osteotomy operation planning

Planning is very critical for jaw osteotomies since there are possible health risks such as infections, relapses, and anastomotic leaks which occur for 10%-20% of the cases (Boléo-Tomé, 1998; Shigeishi et al., 2015). To reduce this risk, a "virtual surgery" is performed on computers using the patient's head scan for planning the surgery (Xia et al., 2000) and observing the possible outcomes. Before the virtual surgery, the patient's 3D digital data of the patient's head is pre-processed: (1) the neck, flesh, nerves are segmented and removed away so that the bones of the skull and mandible remain, (2) optionally, the dental cast is combined with the Cone-beam computed tomography (CBCT) scan to increase the accuracy of the jaw scan, and (3) the volume is smoothed out. Then, the planning stage is carried out using conventional software tools such as Materialise ProPlan CMF 3.0[1]. The planning stage consists

1 `https://www.materialise.com/en/medical/software/proplan-cmf`

of the following three steps:

1. **Resection:** In this step, a part of the jaw is cut and either moved or removed depending on the type of jaw issue. An osteotomy plane is created to indicate a cut.

2. **Mirroring:** In this step, the correct part is mirrored from the other side of the jaw into the damaged part (Brewster et al., 1984). This step is used when a significant amount of the jaw has been cut and needs to be replaced with another bone segment. This step shows the ideal result after the operation (Brewster et al., 1984).

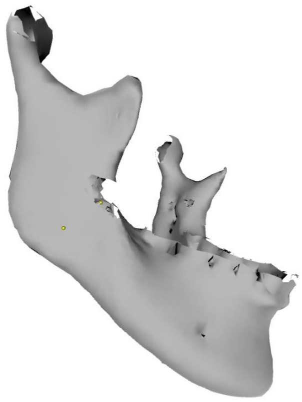
3. **Reconstructing:** In this step, the mirrored part is realigned with the rest of the jaw using a CBCT scan for better reconstruction. This reconstructed structure is used for cutting segments from the fibula bone to fix the jaw. The planning software also allows the user to view how the operation will turn out after the surgery and treatment period.

In this master's thesis, I am focussing on the resection step. In the resection step, parts of the jaw are cut using a saw. The resection step consists of the following three sub-steps:

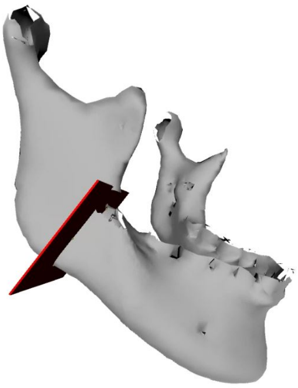
1. **Drawing an osteotomy plane:** The user has to mark a minimum of 3 points on the jaw as shown in Fig. 4.1(a), then an osteotomy plane fitting these points is created as shown in Fig. 4.1(b).

2. **Adjusting the osteotomy plane:** The plane is adjusted so that the thickness of the plane matches the saw blade used in the operation. The cutting area of the osteotomy plane is adjusted so that it goes through the entire or part of the jaw and this is generally checked by viewing the osteotomy plane in the CBCT scan view. The user can also superimpose the nerves on the cut to check whether the osteotomy plane cuts the nerves and adjusts the osteotomy plane if required. The osteotomy plane can be adjusted using translation, rotation, and alignment options, as shown in Fig. 4.1(c) and Fig. 4.1(d).

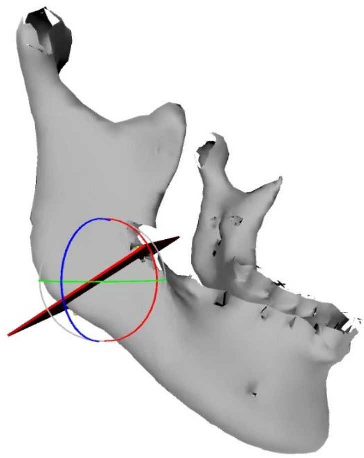
3. **Performing an osteotomy cut using the osteotomy plane:** In the cut operation, the osteotomy plane and jaw are selected and the jaw is cut into two volumes. The user can change the colour of the two parts. This process can be repeated depending on the cuts required. In Fig. 4.1(e), the user cuts a portion of the jaw away by using one osteotomy plane.



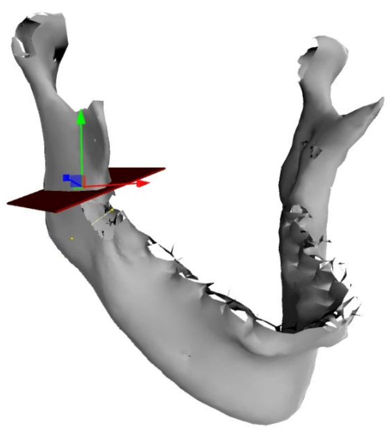
(a) Marking points on the skull to fit a plane



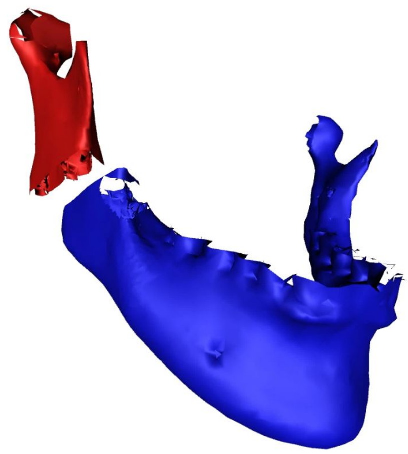
(b) Fitting an osteotomy plane to the marked points



(c) Adjusting the rotation of the plane



(d) Adjusting the translation of the plane



(e) Performing an osteotomy cut using the osteotomy plane

Figure 4.1 *Resection steps of jaw osteotomy operation planning. This 3D jaw model was taken from* `https://free3d.com/3d-model/skull-human-anatomy-82445.html`*. The teeth and upper jaw are removed from the skull model and smoothed into a low poly jaw model.*

4.3 Existing Virtual Reality technologies for jaw osteotomy operation planning

Currently, doctors are performing jaw osteotomy operation planning in hospitals using 2D screen-based GUI with mouse and keyboard. In this current interaction method, the doctors can perform their tasks accurately, but it is a time-consuming process due to a lack of 3D perception. The three-dimensional view of VR technology provides an advantage over 2D screens when it comes to perceiving 3D objects (Olsson et al., 2015; Wagner et al., 1997; Xia et al., 2000) because it is similar to viewing a 3D object in the real physical world (Xia et al., 2000) and it provides "look-around" ability (Olsson et al., 2015).

Existing VR technology-based jaw osteotomy operation planning tools (Hsieh et al., 2002; Olsson et al., 2015; Xia et al., 2000) have been designed with different input devices and interaction techniques for performing the steps of the resection stage to reduce the current execution time from a few days to around 1 hour. These VR environments (Hsieh et al., 2002; Olsson et al., 2015; Xia et al., 2000) used different input devices for resection step: a tracker connected to a tool (a scalpel) (Xia et al., 2000), 3D mouse (Hsieh et al., 2002) and haptics pen (Olsson et al., 2015). They also used different interaction techniques for moving the 3D model: (Olsson et al., 2015; Xia et al., 2000) allowed for the movement of the 3D model in 3D view using direct manipulation, whereas the movement was limited in 2D projections of the 3D model in (Hsieh et al., 2002). There was a difference in how the planes are created and manipulated in these VR environments. In (Olsson et al., 2015; Xia et al., 2000), the osteotomy plane was drawn and adjusted in mid-air using the tool and haptics pen respectively whereas in (Hsieh et al., 2002), the software computed the intersection of the plane swept by the tool with the 3D model.

The input devices explored for 3D object manipulation have certain limitations such as Sensable Phantom pen[2] used in (Olsson et al., 2015) restricts the movements of the user due to the physical extent of the pen arm. (Xia et al., 2000) used a tracker attached to a tool that is like VR controllers, which are the current standard method for manipulating 3D models in VR. Out of hand-based and controller-based methods for object manipulation, users prefer controller because it is more accurate and reliable (Caggianese et al., 2018; Galais et al., 2019; Gusai et al., 2017), but they would like to use hands as they don't need additional hardware (Figueiredo et al., 2018). Similarly, medical professionals want to use their hands as they do not have to use additional hardware and learn how to use it. Xia et al. (2000) proposed the need for a digital glove for holding the scalpel so that the freehand natural hand movements can be captured precisely. With the advancement of AI technology

2 `https://www.3dsystems.com/haptics/`

and computing power, sensors such as Leap Motion[3] can track hand movements in real-time. Thus, hand-based interaction techniques for jaw osteotomy operation planning should be explored further. Out of the three resection steps, maximum accuracy is required for adjusting the osteotomy plane because this is the final step and it compensates for the errors in marking the points. This masters thesis will focus on designing and evaluating hand-based interaction techniques for accurate plane adjustment task.

3 `https://www.ultraleap.com/`

5 Design of Interaction Techniques for Precise Object Manipulation

The design process used in this thesis work is shown in Fig. 5.1. Initially, a contextual inquiry was conducted to understand how the medical professionals adjusted the plane while performing the jaw osteotomy operation planning. This would help in defining the task of osteotomy plane adjustment in the jaw osteotomy operation planning process. In addition, this would help in designing appropriate interaction techniques for precisely manipulating the osteotomy plane. Next, literature review of the existing hand based virtual reality object manipulation interaction techniques was conducted to understand the various design factors used by existing research and the results of the evaluation of these interaction techniques. After this, potential design factors for interaction techniques were selected for our task. Initial interaction techniques were designed and implemented based on these factors. Pilot tests with 2 HCI researchers was conducted to get feedback on these designs. The interaction techniques were refined based on the feedback.

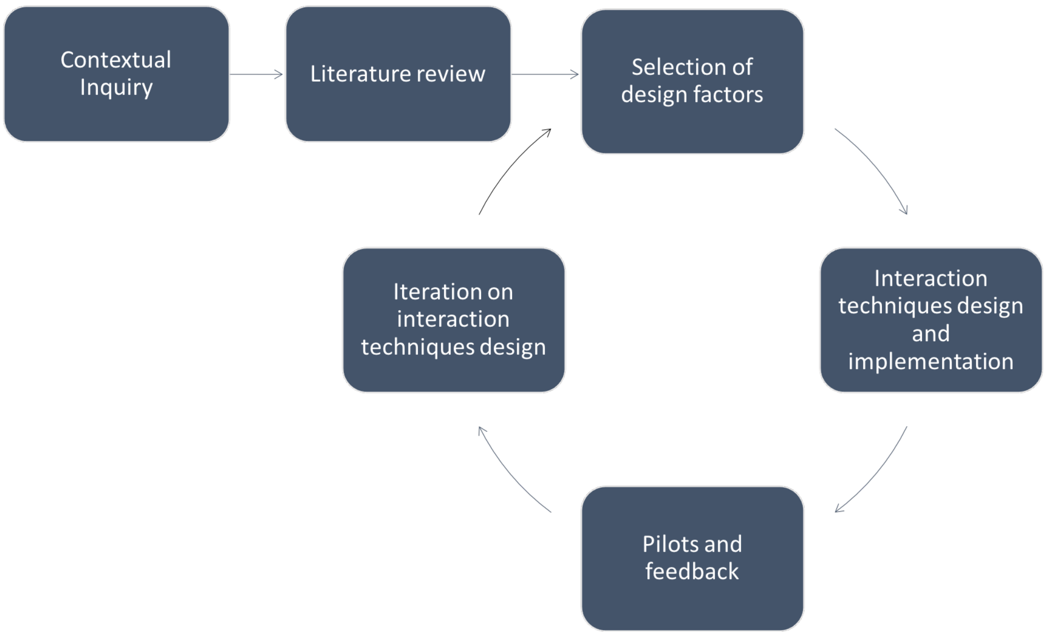


Figure 5.1 *The design process.*

5.1 Contextual Inquiry

One radiologist with jaw osteotomy operation planning experience was recruited for the contextual inquiry. The contextual inquiry was carried out in two phases. The first contextual inquiry was conducted in the hospital. In the first phase, the participant was asked to explain the jaw osteotomy operation planning process. The second phase was conducted online through online video calling software. The participant was present in the lab and was asked to shared the computer screen. The participant was asked to show how the jaw osteotomy operation planning process would be performed for real patient data using software of choice. The moderator asked follow up questions to clarify details about the task. The contextual inquiry took around four hours in total.

5.1.1 Observations

The following observations were made regarding the osteotomy plane adjustment task. This was useful for creating a task for the study and designing the interaction techniques.

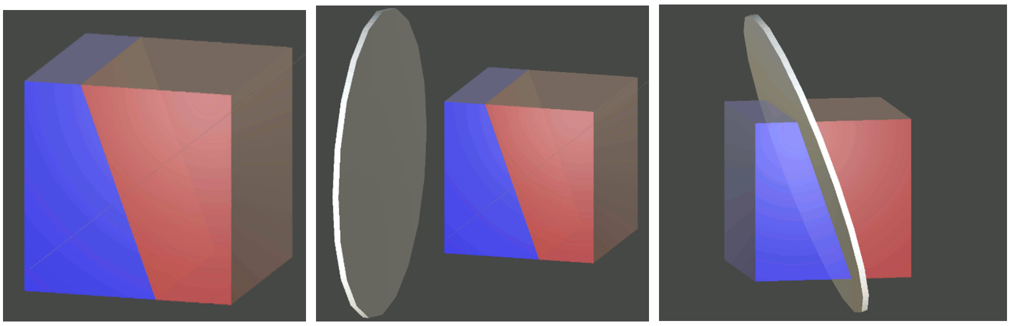
Based on the contextual inquiry, medical professionals performed osteotomy plane adjustment based on two situations. The first situation is to exclude an area, for example, excluding a cancer affected area where a fragment of the bone has to be removed. The second situation is to place the osteotomy plane between specific points and angling between them, for example, to position the osteotomy plane between two anatomical landmarks such as two teeth and angled based on the alignment of the teeth.

In both situations, the medical professional would want to place the plane accurately between certain areas. In first situation, the decision is taken to exclude certain area and in the second case, the professional would want to consider anatomical points to place the cutting plane so that it does not cut through unevenly or considering logical cuts such as between teeth rather than through a teeth. The medical professional would want to have as minimal damage as possible.

For designing the task and interaction techniques, both situations would be considered.

5.1.2 Task for precise object manipulation

Based on the observations from the contextual inquiry, the task for the interaction and study was created. In this task, there is a cube which is precut into two coloured segments: red and blue as shown in Fig. 5.2(a). A plane is placed close to the cube as shown in Fig. 5.2(b).The task was to align the plane in order to separate the two coloured segments of the cube as shown in Fig. 5.2(c). The participant can only interact with the plane and not the cube. Similar to how the medical professionals decide to place the cutting plane to exclude an area or between the anatomical points, here, the user must place the plane between two shaded regions and try to align to the edges of the cube cut by the plane.



(a) Cube (b) Start state (c) End state

Figure 5.2 *The precise manipulation task for the study*

5.2 Designing interaction techniques

Based on the above design decisions, the following two interaction techniques were designed. The descriptions of these interaction techniques are discussed below in detail.

5.2.1 Interaction techniques design decisions

Previous interaction techniques for object manipulation in virtual reality explored various design factors such as gestures, metaphors, widgets, number of hands, C/D gain, constraints, feedback as listed in Fig. 3.4. In addition, to these design factors, there are several human factors that should also be considered while designing interaction techniques for object manipulation. These factors are expectation of physics and collision, spatial understanding and viewing angle. Out of these human factors, physics and viewing angle were explored for designing interaction techniques in this thesis work.

Gestures such as pinch and grasp have been already used for virtual reality but these are not appropriate for virtual reality as the start and end of the gestures may not be detected at the proper time. Thus, dynamic gestures which do not require the start and end phase to be detected would be more appropriate. Therefore, it would be better if hand gesture recognition is not required. Expectation of collision is an important factor when people interact with objects. Collision in virtual reality requires only hand tracking and this eliminates the problem of gesture recognition completly. Fig. 3.5 shows the gestures used by humans while interacting with objects in real world. Using such gestures would make it familiar for the users to learn and use. Out of these gestures, the hand gestures with no clenching and grasping (no-prehension) or arm based gestures would be appropriate while colliding with objects in virtual space. Through pilots with two HCI researchers, push and poke were selected to be appropriate for colliding with objects in virtual reality with push being appropriate for long distances and poke for small distances.

The view of the virtual world is based on the current viewing angle. The user has to move around to see other parts of the world and perform operations in those parts. The operations done in the vicinity of the current position are more accurate than parts that are occluded. Based on this, a widget with a custom axis based on the head pose of the user was created. This interaction techniques allows the user to view appropriate translation and rotation axes based on the current head pose of the user thus making give most accurate axes options possible. To make the manipulation more granular, C/D gain was added to scale down the movements.

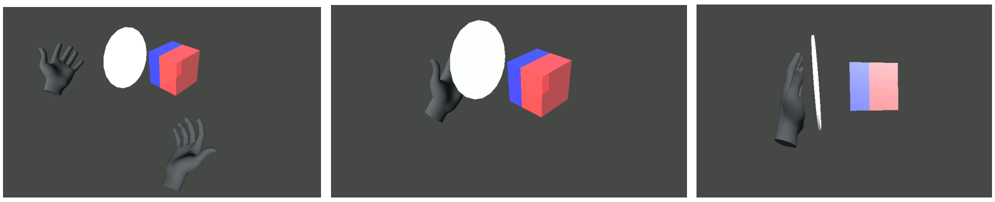
The details of these interaction techniques are explained below.

5.2.2 Push and poke

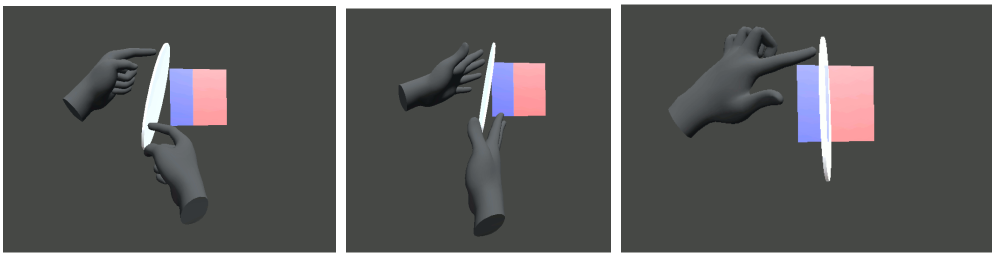
Push and poke interaction is a collision based direct manipulation approaches. The user has to use bare hands as shown in Fig. 5.3(a). The user can use either palm or fingers to collide with the plane to translate or rotate the plane, for example, in Fig. 5.3(c), the user is using the palm to push and translate the plane forward and in Fig. 5.3(d), the user is poking the plane with two fingers to rotate the plane. In Fig. 5.3(f), the user can use the same poke gesture to nudge the plane slightly to correct the position and orientation of the plane. In this interaction technique, the system works obeying the laws of physics in zero gravity. The plane is a rigid body with a mass of 47.95, drag of 19.6 and angular drag of 29.76. These values are initially set using trial and error and later validated with initial pilots with 2 HCI researchers.

5.2.3 Viewing angle with C/D gain

This interaction technique uses head pose for determining the custom axis and C/D gain for making the movements granular. In this method, as shown in Fig. 5.4, there are two concentric rings which appear around the object and the rings always facing the user head (camera). The rings follow the position of camera. Both the rings have a ball which also follows the hand position. The inner ball which is blue coloured is used for translation and the outer ball which is red is used for rotation. Based on how the user grabs the rotation or translation ball, an axis is created perpendicular to the plane and in the direction of the viewing angle when the ball is grabbed. The user can pull along the axis as shown in Fig. 5.4(b) and Fig. 5.4(d). Once pinched the ring will freeze until it is released. The distance between the initial pinch location and current pinched finger position is scaled down using C/D gain factor of 0.1X to reduce the motion in translation and rotation operations. This interaction stops when the pinch is released.

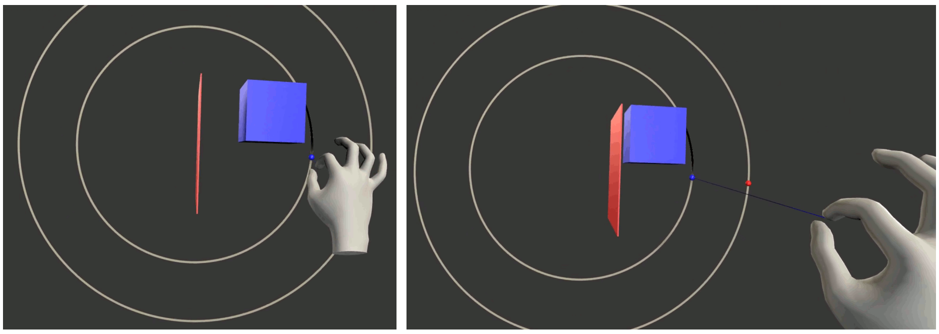


(a) Initial state (b) Push with palm (c) Push with palm

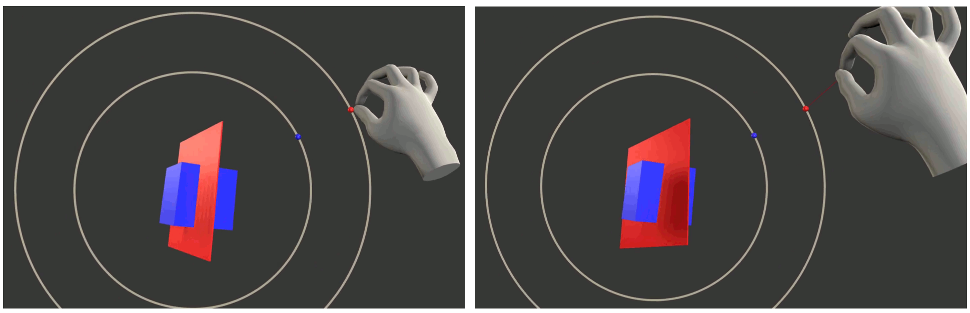


(d) Poke with two fingers (e) Combination of push and poke with two hands (f) Poke with one finger

Figure 5.3 *Examples of push and poke interaction to translate, rotate and nudge a plane*



(a) Grab the translation handle (b) Pulling the translation handle to translate the plane



(c) Grabbing the rotation handle (d) Pulling the rotation handle to rotate the plane around the center of the plane

Figure 5.4 *Examples of viewing angle with C/D gain interaction technique to translate, rotate a plane*

5.2.4 Pilots and feedback

Pilots of the interaction techniques with the task, mentioned in Sec. 5.1.2, was conducted with 2 HCI researchers to understand whether these interaction techniques work and how they can be improved.

No negative feedback was given for the push and poke interaction. Based on these pilots, no changes were made to this interaction technique.

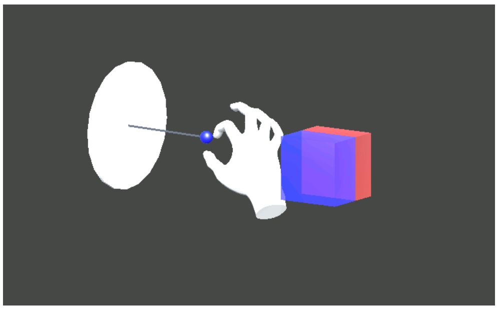
Some issues were discovered with the viewing angle with C/D gain. The participants found it hard to manipulate the plane based on viewing angle as they constantly had to move around the object. Both participants wanted to manipulate the plane based on the position of the hand rather than the head pose. The participants felt that the circle axis was redundant as it did not signify any meaning to them. One participant found rotating the plane around the center very difficult because "I need to first think about which axis and then think about the rotation angle, which is not natural". This participant suggested that this interaction technique should support "moving this point of the plane to a another point and maybe other side of the plane remains fixed".

5.2.5 Custom axis with C/D gain

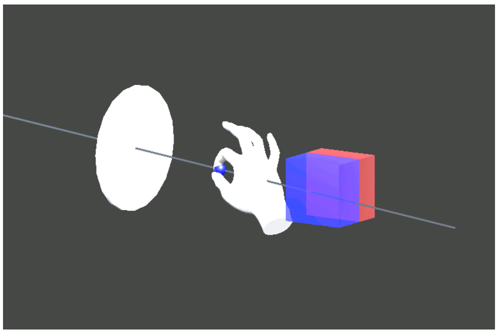
Based on the feedback from the pilot, changes were made to the viewing angle with C/D gain.

The concentric rings were removed. Instead of viewing angle (head position), translation and rotation handles were created based on hand position. The translation handle was created at a particular distance from the plane and in the direction of the hand as shown in Fig. 5.5(a). The rotation handle was created on the plane at the position of dominant hand's index finger position as shown in Fig. 5.5(d). The user had to grab the handles and move to translate as shown in Fig. 5.5(b) and rotate as shown in Fig. 5.5(e). Once the handle was grabbed, the handles stop following the hands and a guiding axis was created. The user drags the handle in the direction of the guiding axis like a slider to translate and rotate in the desired direction. The translation and rotation interaction stop when the user releases the pinch. C/D gain factor of 0.1X was applied to scale down the motion in translation as shown in Fig. 5.5(c) and rotation as shown in Fig. 5.5(f). For rotation, instead of rotating around the center, a custom pivot point on the opposite side of the plane was created as shown in Fig. 5.6. This makes the rotation operation work similar to the nudge operation in push and poke interaction.

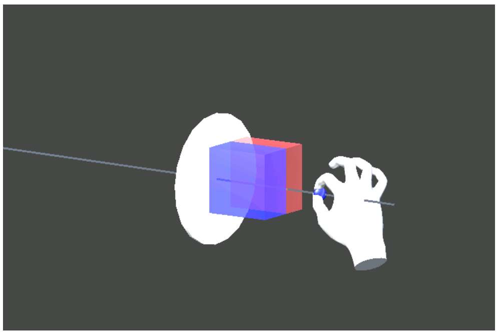
Based on the design changes, the name of this technique was changed to "custom axis with C/D gain".



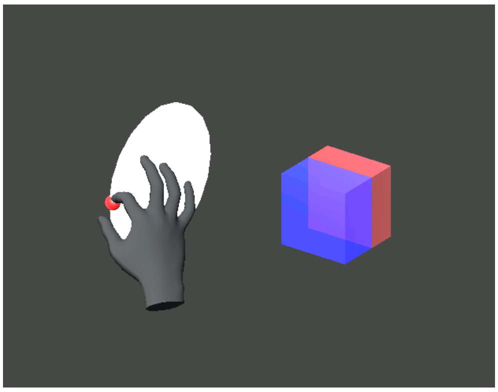
(a) The translation handle appears based on the hand position



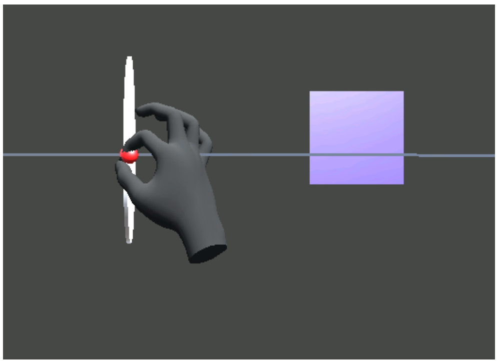
(b) Pulling the translation handle to translate the plane, a guiding axis is shown



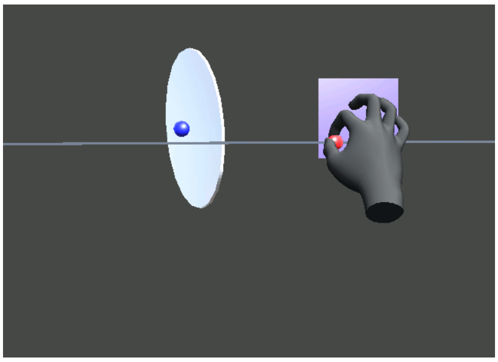
(c) The plane is translated based on the movement and C/D gain



(d) The rotation handle appears on the plane based on closest index finger position



(e) Pulling the rotation handle to rotate the plane



(f) Pulling the rotation handle to rotate the plane around the center of the plane

***Figure* 5.5** *Examples of custom axis with C/D gain interaction technique to translate, rotate a plane*

5.2.6 Implementation

The interaction techniques were implemented with Unity3D software. Oculus Quest headset (64GB memory) was used for deploying these interaction techniques. This headset uses four onboard optical cameras for performing hand tracking. Oculus intergration was used to build the hand tracking functionality. It was provided by

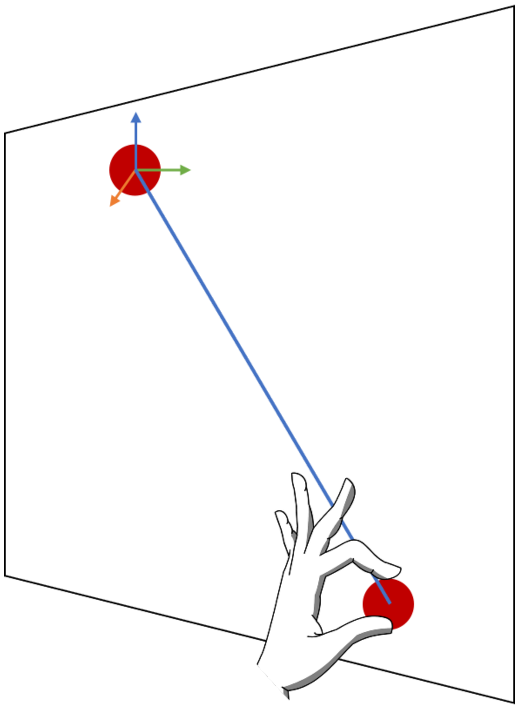


Figure 5.6 *The position of the custom pivot point in rotation in custom axis with C/D gain.*

Oculus (Oculus Quest, 2020) in unity asset store.

6 Experiment

To evaluate these designs for precise manipulation, a controlled experiment with 12 participants was conducted. The conditions were presented in counterbalanced order using Balanced Latin Square to reduce learning effects. In this experiment, pinch based direct manipulation as shown in Fig. 6.1 is used as baseline.

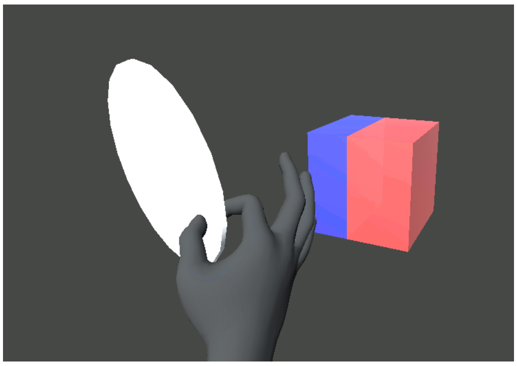


Figure 6.1 *The pinch based direct object manipulation interaction technique.*

In this experiment, task completion, precision, perceived ease of use, learnability, confidence, intuitiveness, naturalness, hand tiredness were measured.

6.1 Research questions

This experiment aims to answer the following research questions:

RQ1: Which interaction technique is more accurate and preferred for object manipulation in virtual reality?

An interaction technique should provide both accuracy for the task and should be suitable for users to use. This research question investigates which interaction is suitable for both task and the user and whether there is any trade off.

RQ2: Do these interaction techniques support both small and large movements and which one(s) is required for precise object manipulation?

According to Turner and Van De Walle (2006), there are two types of movements for object selection task, fast and imprecise movements as well as slow and precise movements. In the study conducted by Mendes et al. (2016), they found that a design that helps with fine grain adjustments while preventing unwanted movements helped the participants to achieve higher precision than direct manipulation. Here, we understand the differences between the interaction technique usage in terms of movements and understand whether there is a need for one type of movement over the other.

RQ3: Should hand-based interaction techniques use gesture recognition?

Hand tracking is challenging due to the technology. Occlusion of hands, egomotion of the head, variations in hand size and shape (Sharp et al., 2015), lighting, noise in the sensor (Oculus Quest, 2020) are some of the reason for hand tracking to be unreliable. Grab and pinch were the most commonly explored gesture that is compared with other input devices (Caggianese et al., 2018; Galais et al., 2019; Gusai et al., 2017). Hand is a versatile input medium and hand is capable of using much more gestures and interaction methods than just Grab and pinch (Feix et al., 2015).

Push and poke interaction technique is designed to not use gesture recognition. Whereas pinch uses gesture recognition for grab and release and custom axis with C/D gain uses gesture recognition for grab and release for translation and rotation handles. By comparing these interaction techniques, we understand whether it is better to design hand based interaction techniques with or without gesture recognition and how does this decision impact the performance and experience.

6.2 Participants

Twelve participants (7 male, 5 female) were recruited for this study using snowball sampling. The age was from 24 years to 38 years with a mean of 30 years and standard deviation of ±5. Nine participants were university students and three participant were full time employee. We also asked about their experience in using virtual reality devices and experience of using their own hand as input to use the virtual reality applications. It was rated between 0 to 7, where 0 means no experience, 1-2 means novice users experience, 3-5 means medium experienced users and 6-7 means expert users. The participants that we recruited has experience in using virtual reality and using hands in virtual reality as shown in Tab. 6.1. The handedness of the participants was also collected. The handedness information was used to select the dominant hand for the custom axis with C/D gain interaction technique.

	Using Virtual Reality	Using hand as input in Virtual Reality
No experience	1	4
Novice	3	3
Medium	5	3
Experts	3	2

Table 6.1 *Participants experience level*

6.3 Experiment design

In this, experiment we compared the two proposed designs with the baseline approach of direct manipulation. Thus, the experiment had a total of three conditions:

1. Condition 1: Direct manipulation (baseline)
2. Condition 2: Push and poke based collision in zero gravity
3. Condition 3: Custom axis with C/D gain

Within-subject evaluation was used to compare the different conditions with the same participants. The conditions were presented in counter balanced based on Balanced Latin Square order to reduce the effect of ordering the conditions when analysing the results.

6.3.1 Task

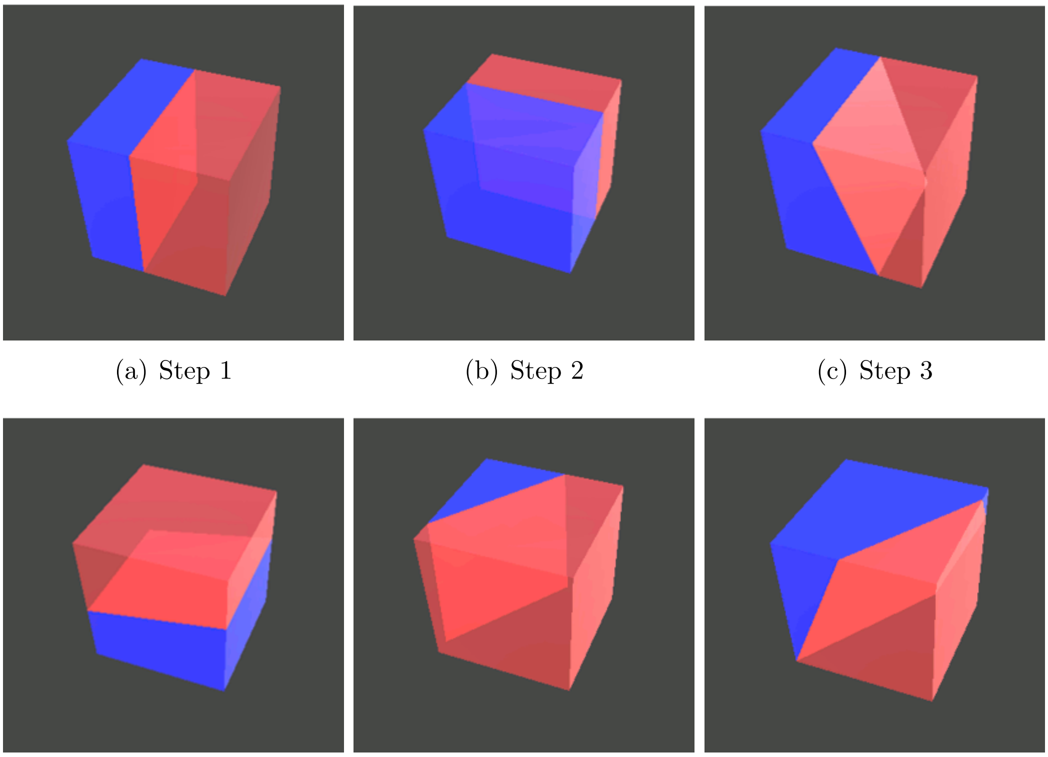
The task for precise plane alignment is discussed in Sec. 5.1.2.

A total of six planes are generated and presented to the participant under each condition. One plane is shown at a time. A plane is placed close to the cube for the first task as shown in Fig. 5.2(b) and kept in the previous position for the next tasks. The first plane position and the precut plane positions are generated such that the user had to perform minimal translation and rotation. The six precut planes used in the three conditions are shown in Fig. 6.2.

Since this task is of precise manipulation, the plane has to be close enough to the ground truth precut plane in order for the task to be marked as complete by the participant. Here, the accuracy is calculated by the Eqn. 6.1.

$$accuracy = \frac{distance\ accuracy + 2 \times angle\ accuracy}{3} \tag{6.1}$$

where the distance accuracy and angle accuracy are defined in Eqn. 6.2 and Eqn. 6.3 respectively. The distance in distance accuracy Eqn 6.2 is computed as the distance of the plane to the closest point on precut plane. Then distance is thresholded at 35cm and transformed into percentage from 0 to 100% to compute the distance accuracy. The distance in angle accuracy in Eqn. 6.3 is computed as the Euclidean distance between the normal of precut plane and the normal of the plane; transforming the x, y, z components of these planes into absolute values before computing the distance so that the normals end up in the same quadrant in the 3D coordinate space. The distance is then thresholded at 1 and transformed into a percentage from 0 to 100% to compute the angle accuracy.



(a) Step 1 (b) Step 2 (c) Step 3

(d) Step 4 (e) Step 5 (f) Step 6

Figure 6.2 *The six planes used for the precise manipulation task in the study.*

$$distance\ accuracy = \begin{cases} 0 & distance > 0.35 \\ 100 - \frac{distance*100}{0.35} & otherwise \end{cases} \tag{6.2}$$

$$angle\ accuracy = \begin{cases} 0 & distance > 1 \\ 100 - distance \times 100 & otherwise \end{cases} \tag{6.3}$$

The threshold for this accuracy is kept at 95%, thus the participant has to be at least 95% accurate to finish the task and consider it as complete.

6.3.2 Measures

The following measures were logged and collected to evaluate the interaction techniques.

1. Task completion time (seconds)

2. Number of interactions with the plane to complete the task.

3. Accuracy of the plane alignment as per the Eqn. 6.1 when the task is completed.

In addition, the participants were asked about the subjective perception for the following measures on a Likert scale from 1 to 7. These measures were adapted from SUS (“Sus: a “quick and dirty’usability”, 1996).

1. Learnability
2. Ease of use
3. Confidence
4. Intuitive
5. Naturalness
6. Perceived precision
7. Self reported hand tiredness
8. Using daily

6.4 Experimental Setup

The headset was connected to the laptop using a USB-C cable. The participant was asked to do the study standing up so that they can easily move around the 3D space to observe the cube from different angles and perform the precise manipulation task as shown in Fig. 6.3.

The Unity3D environment contains the cube, the plane and a dashboard with the accuracy and task number and a button for progressing the task as shown in Fig. 6.4.

The button in Fig. 6.4 is initially grey. The button becomes red when the threshold for accuracy is reached. Participant has to press the button to advance to the next task in the condition.

6.5 Procedure

6.5.1 Introduction and Background Information Collection

The participants were initially welcomed. The purpose of the study was explained. The task was explained and hand based interaction was mentioned. The participants were made aware that the aim of the study is accuracy and not speed so that they focus on accuracy. The data to be collected was explained. The participants were informed that they could discontinue the study whenever they wanted for any reason including sickness in VR. The participants were asked to sign the consent form. After this, information such as age, gender, profession, dominant hand, experience in using

Figure 6.3 *A participant performing the precise manipulation task wearing a VR headset and standing up.*

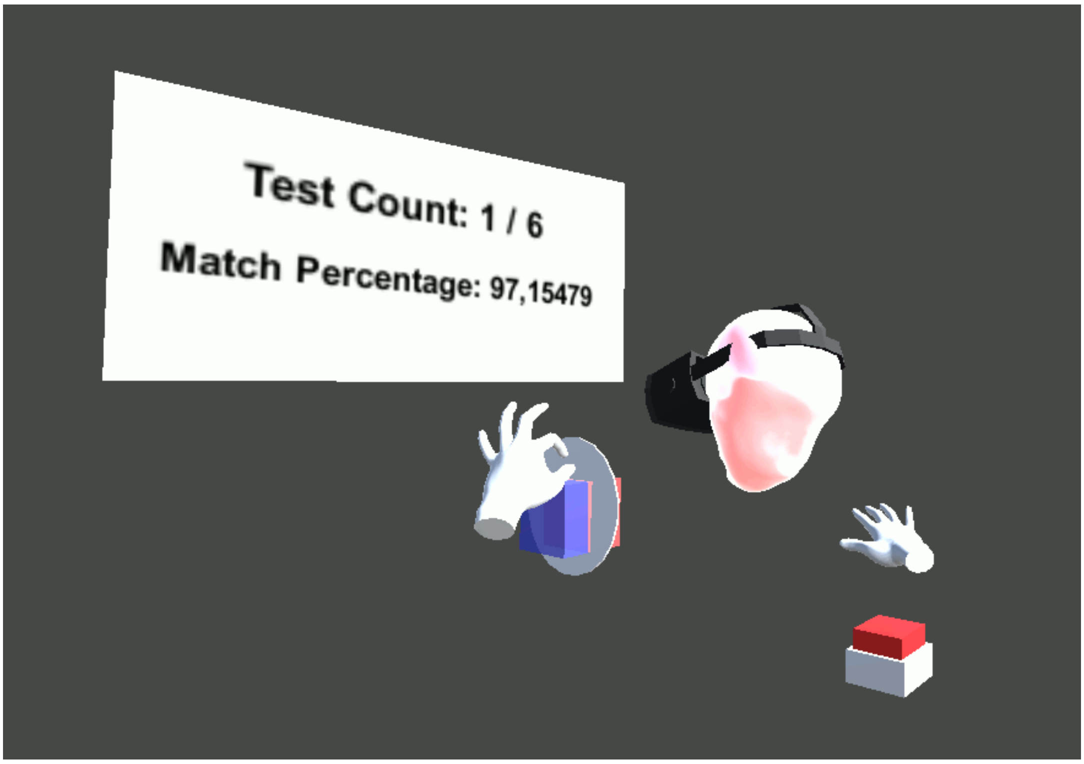


Figure 6.4 *The experimental setup of the precise manipulation task in Unity3D environment.*

virtual reality and interacting with hands in virtual reality (Likert scales from 1 to 7) were collected from the participants. The background questions are included in Appendix A.

6.5.2 Training task

A training task was used to help the participant understand how to use the interaction technique and get familiar with the task. The training task was similar to the real task, however, the precut plane was randomly generated.

First, the moderator explained the interaction technique by putting the headset, sharing the VR view on the laptop screen and explaining the interaction technique. Then the headset was put on the participant and the participant tried to use the interaction technique. The moderator verbally guided the participant in understanding the interaction technique. The participant could ask the moderator questions on how to use the technique. The participant could take as much time as needed to practice the task. The participant could press the button and a new precut plane would be randomly generated. The participant could continue with the study when they felt confident.

6.5.3 Study

In the actual study, the participant used the interaction technique according to the condition to move the plane to get a minimum accuracy of 95%. The participant could move on to the next task in the condition, when the button turned red (achieved a minimum accuracy of 95%). The participant had to move six planes in one condition.

6.5.4 Survey and Semi-structured Interview

After each condition, participants were asked to fill in a survey with the subjective perception questions. In addition, participants were asked follow-up questions for any subjective rating with a value below 4. The participants were asked about the positive(s) and negative(s) of the interaction techniques. The questionnaire used in the subjective condition evaluation is included in the Appendix B.

After the three conditions were finished, the participants were asked to rank the three conditions based on their preference for most liked, most precise, best suited for novice user and most potential to be developed further. The questionnaire used in the subjective post experiment evaluation is included in the Appendix C.

6.6 Analysis

One-way ANOVA analysis was performed to compare the effect of the condition on the quantitative data to understand if there is an effect of condition on task completion time, precision, number of interactions subjective ease of use, learnability, confidence, hand tiredness, naturalness, intuitiveness, confidence, precision. To understand if there were any significant difference between conditions, posthoc pairwise comparison was performed using Bonferroni Correction with a corrected p-value of $0.05/3 = 0.0167$. If there is no statistical difference between all the pairs of conditions for objective accuracy and subjective precision, Friedman Chi Square test was used to check if there is a statistical difference in terms of ranks.

7 Results

In this chapter, the results of the study are presented. The conditions in the study are compared in terms of objective measures of accuracy, task completion time and number of interactions as well as subjective measures of learnability, usability, naturalness, confidence, precision, preference and most potential for future development. The comparison of subjective measures is further explained with quotes from the participants collected in the semi-structured interview of the study.

7.1 Objective measures

Three objective measures of accuracy, task completion time and number of interactions were collected. The following discusses the statistical differences between the conditions in terms of these objective measures.

7.1.1 Accuracy

The distribution of the accuracy of the final plane placement for each condition is shown in Fig. 7.1. There are no significant differences between the conditions in terms of accuracy. However, there is a statistical difference in the ranks of the median accuracy across the tasks between all the conditions using Friedman Chi Square (p=0.002). The average rank of push and poke is 1.25, the average rank of custom axis with C/D gain is 2.08 and the average rank of pinch is 2.67.

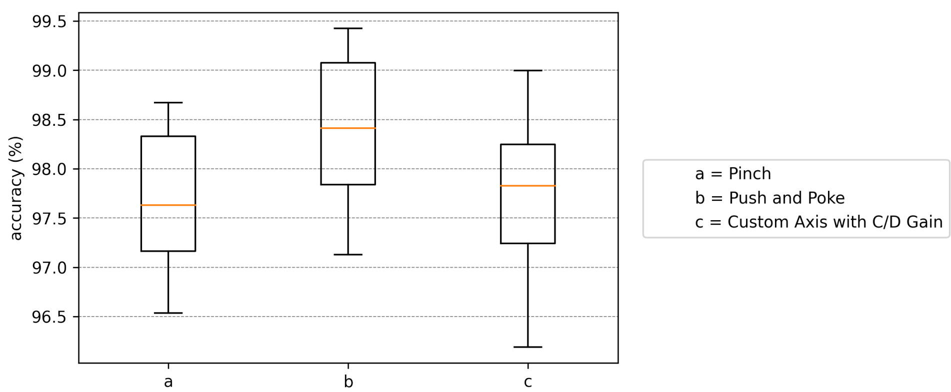


Figure 7.1 *The distribution of the accuracy of the final plane placement (the median value across tasks in a condition) for each condition.*

7.1.2 Task Completion Time

The distribution of the accuracy of the task completion time for each condition is shown in Fig. 7.2. The time to complete the task using custom axis with C/D gain is significantly much higher than both pinch (p=0.0001) and push and poke (p=0.0028).

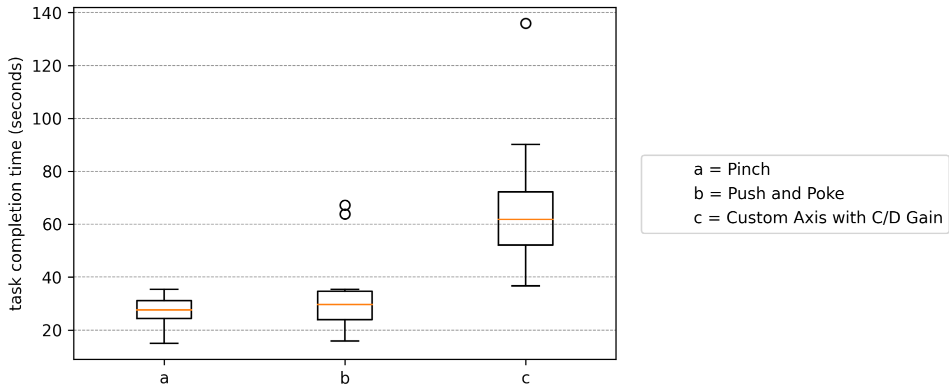


Figure 7.2 *The distribution of the time taken to complete the task (the median value across tasks in a condition) for each condition.*

According to the the seven stages of action proposed by Norman (2013), interaction consists of evaluation stage and execution stage. Based on this, the evaluation and execution times are computed. Execution time is calculated as the time between the moment an interaction starts and when it finishes, for example, when the plane is pinched and when it is released. Evaluation time is calculated as the time between the moment an interaction finishes and the next interaction starts. Fig. 7.3 shows the mean of the mean evaluation time in a task across the trials in different conditions and Fig. 7.4 shows the mean of the mean execution time in a task across the trials in different conditions. In the case of evaluation time needed for each interaction, push and poke required significantly much less time than pinch (p=0.0) and custom axis with C/D gain (p=0.0001). However, for the execution time needed for each interaction, push and poke required significantly less time than only pinch (0.0124). This shows that the user had to use significantly less decision making and less execution time in push and poke condition than both pinch and custom axis with C/D gain to achieve similar levels of incremental improvement in accuracy.

7.1.3 Interactions in a task

The distribution of the number of interactions in a task for each condition is shown in Fig. 7.5. The number of interactions in a task when using pinch is significantly

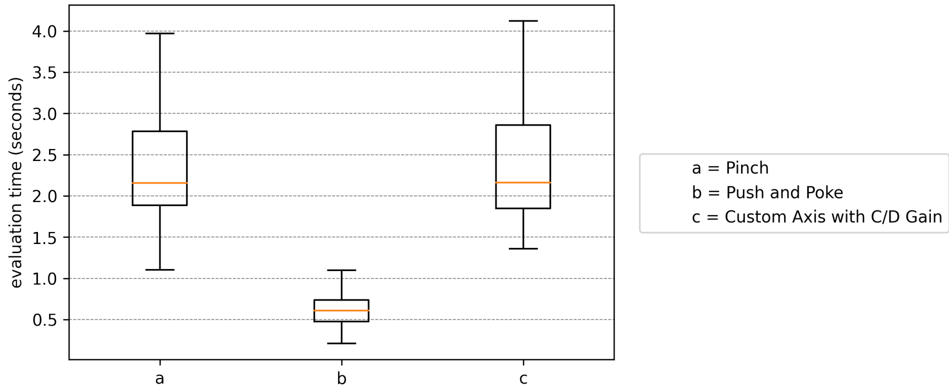


Figure 7.3 *The distribution of the mean of the mean evaluation time in a task across the trials in different conditions.*

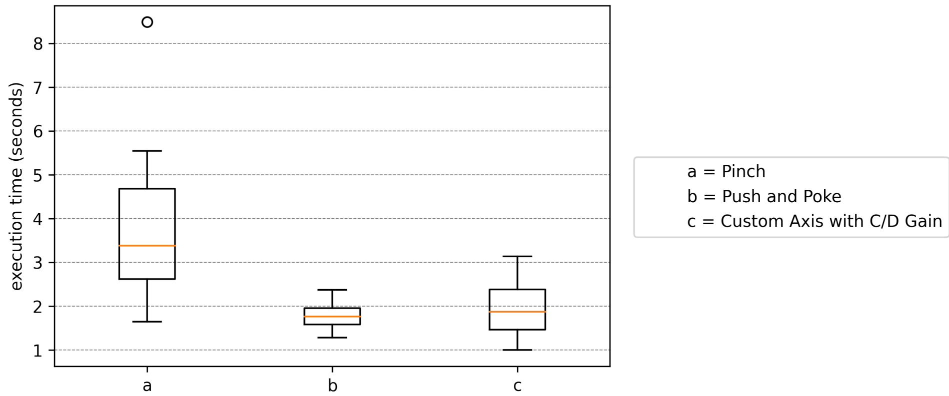


Figure 7.4 *The distribution of the mean of the mean execution time in a task across the trials in different conditions.*

much lower than both push and poke (p=0.0001) and custom axis with C/D gain (p=0.)

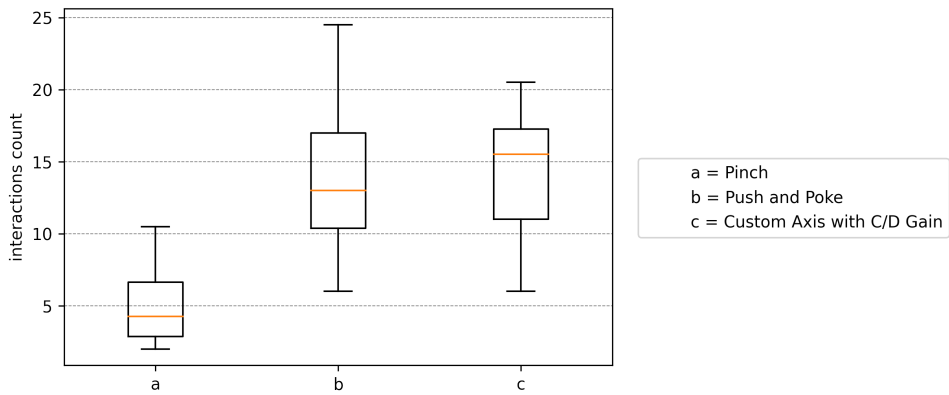


Figure 7.5 *The distribution of the number of interactions in a task (the median value across tasks in a condition) for each condition.*

According to Graham and MacKenzie (1996), in object selection task the user movement can be divided in to two phases: (1) initial fast and imprecise movement to the target and (2) final slow and precise movements before touching the object (Graham & MacKenzie, 1996). To understand when and whether fast and slow movements are made during the task, trend lines of the accuracy changes with interaction timestamp in task are plotted. Fig. 7.6 shows the accuracy trend for each participant and across all participants for each of the interaction techniques. The accuracy over 70% is considered for calculating the trend, as the starting accuracy for different trial is different. In addition, each task took different times to complete, so the task time is normalised to the median of the median time for task completion time for that interaction technique and then the trend line is fitted. The trend line is fitted using Support Vector Regression (SVR) method with a radial basis function (RBF) kernel with C = 10, gamma = 0.1 and epsilon = 0.1. Fig. 7.7 shows the trend of the accuracy changes across the participants for each interaction technique. These trend lines are computed in the similar way with the same parameters.

To further understand the minimum and maximum movements offered by these interaction techniques, the distribution of the median of the minimum accuracy changes across participants for each interaction technique are plotted. Fig. 7.8 shows the distribution of the median of the minimum accuracy changes across participants for each interaction technique. The consecutive accuracy changes more than 0.01 were considered in this calculation as small changes might be created using noisy interactions. The ranges of minimum movement in push and poke and custom axis with C/D gain for the participants is much less than the pinch. The small movements made in custom axis with C/D gain is significantly smaller than only push and poke (p=0.0006). Fig. 7.9 shows the distribution of the median of the maximum accuracy changes across participants for each interaction technique. The large movements made by pinch interaction is much smaller than both custom axis with C/D gain (p=0.0088) and push and poke (p=0.017).

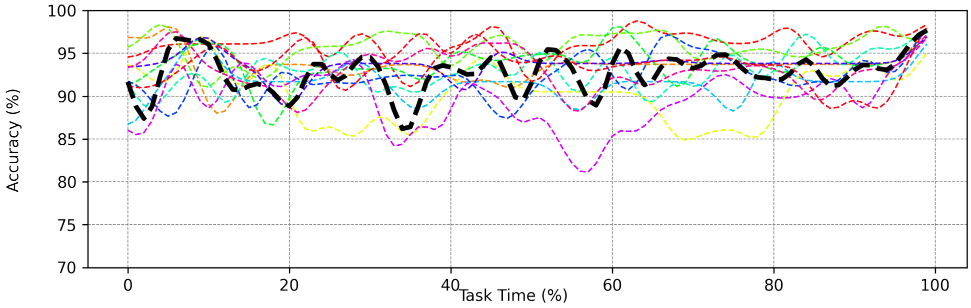
7.2 Subjective Evaluation

The distribution of the subjective evaluations values for each condition is shown in Fig. 7.10. There was no statistical significance in terms of hand tiredness between the three conditions.

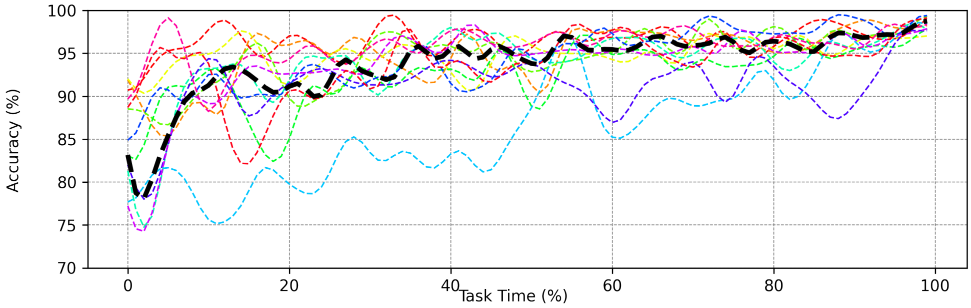
7.2.1 Learnability

Participants felt that custom axis with C/D gain was significantly harder to learn than push and poke (p=0.0005).

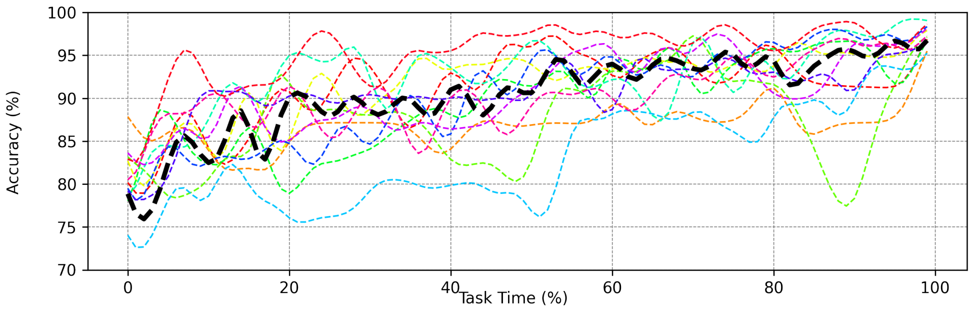
Participants felt that the familiarity of pinch, push and poke gestures in real life



(a) pinch



(b) push and poke



(c) custom axis with C/D gain

Figure 7.6 *The trend of accuracy for each participant and across all participants for each of the interaction techniques.*

would help novices in learning these interaction techniques. P12 said that "pinch seemed like something I do in real life" and P5 said that, "pinch and push are familiar from real world". P4 said that, "Push and poke are daily based gestures that can be easy for a novice user to work with".

Participants felt that the custom axis with C/D gain had a learning curve. P7 said that, "I think it has a learning curve", P8 said that, "it was challenging at first". However, some participants felt that rotation handle was easy to learn and use. Some participants wanted more time to learn, P6 said that, "need more time to practise". Some participants felt that they need to learn using handles. P8 said that,

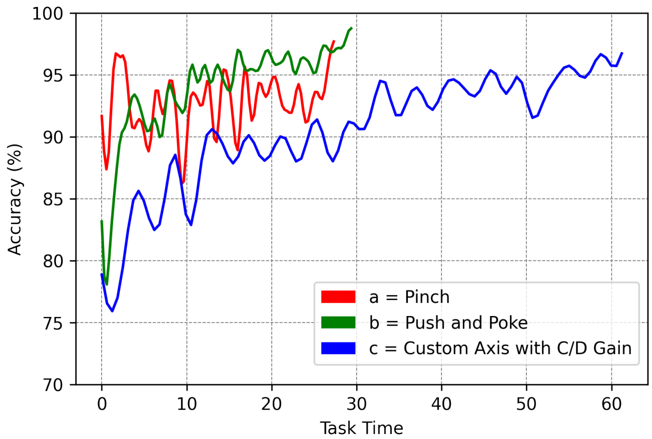


Figure 7.7 *The comparison of the trend of accuracy across all participants for each of the interaction techniques.*

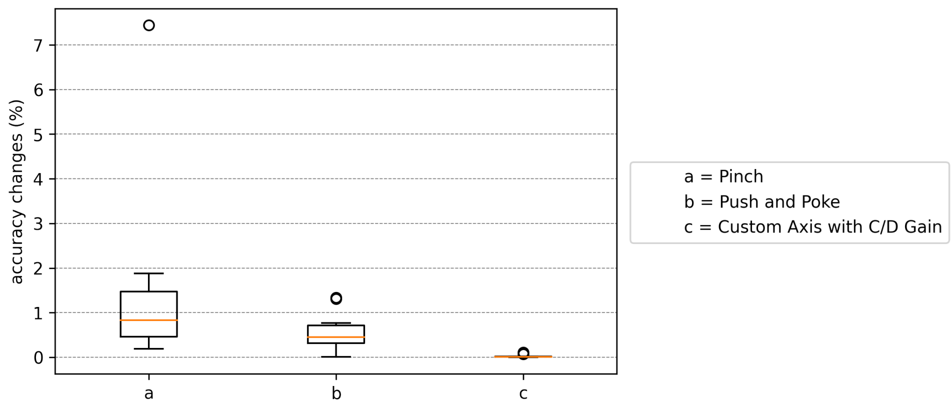


Figure 7.8 *The distribution of the median of minimum accuracy changes across participants for each interaction technique.*

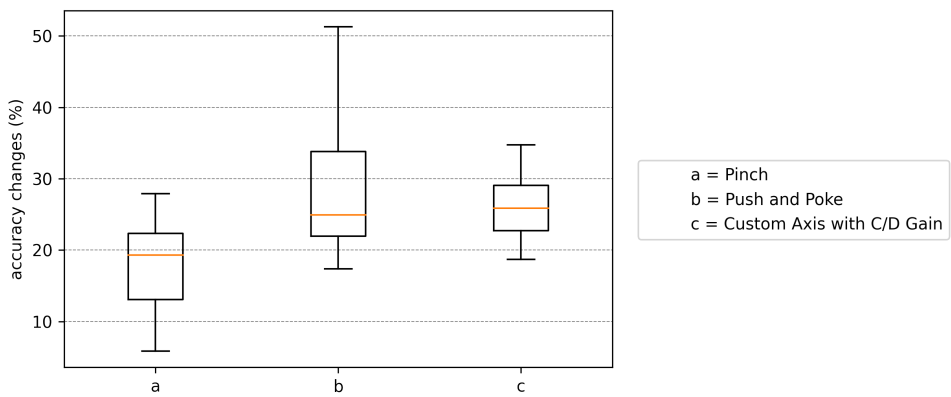


Figure 7.9 *The distribution of the median of maximum accuracy changes across participants for each interaction technique.*

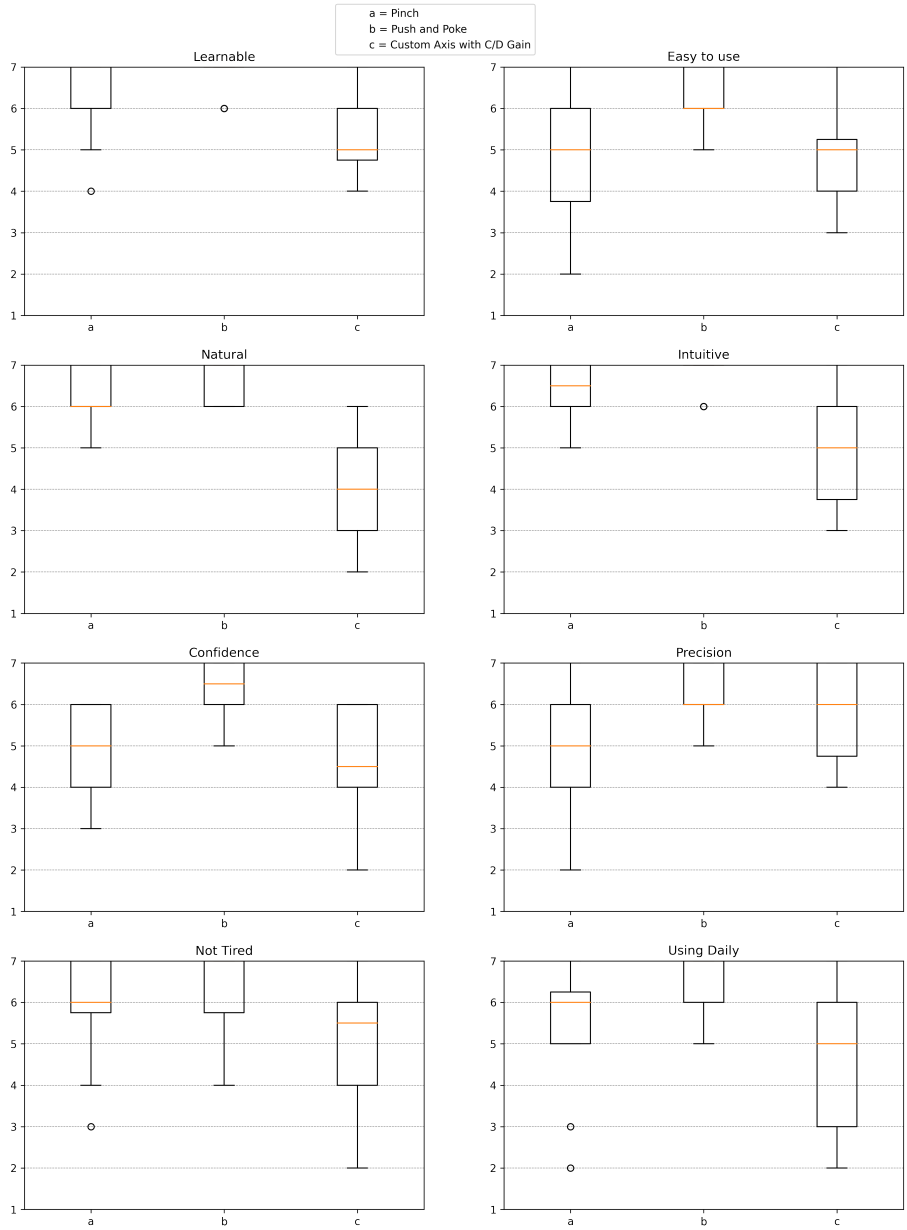


Figure 7.10 *The distribution of subjective evaluations values for each condition.*

"the use of the handles was some what different and not so familiar to me". P12 said that, "grabbing blue ball is hard". Some participants felt that C/D gain feature had to be learnt. P11 said that, "There's some rubberband effect which might take some time to get used to and it might affect the accuracy of the manipulation".

When asked to rank the conditions in terms of suitability for novice users, participants felt that push and poke was the most easiest to learn, followed by pinch and then custom axis with C/D gain as mentioned in Tab. 7.1. This ranking is related to learnability of the techniques. P10, P12 mentioned that pinch and push and poke are more suitable for novice users as these are “natural” and “easiest to learn”. P6 mentioned that custom axis with C/D gain “is hard to learn”.

Condition	Ranking		
	1st	2nd	3rd
Pinch	4	7	1
Push and poke	8	4	0
Custom axis with C/D gain	0	1	11

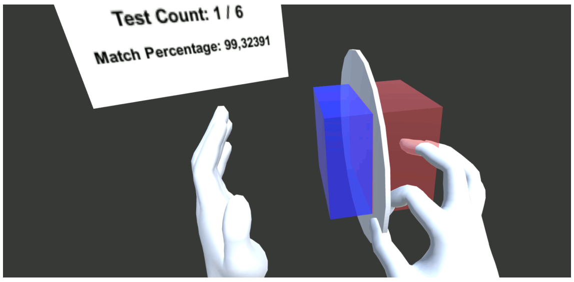
Table 7.1 *Accumulated count of users preference based on their ranking for which interaction technique they feel is best suited for novice user. One is the most appropriate and three is least appropriate.*

7.2.2 Ease of use

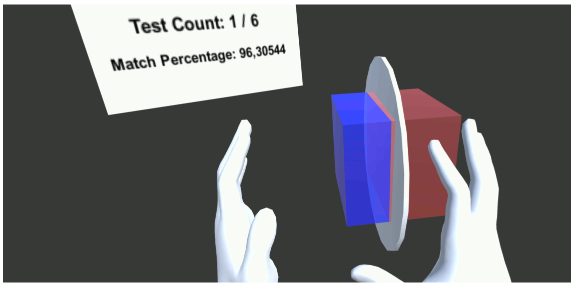
Participants felt it was significantly more easy to use push and poke than custom axis with C/D gain (p=0.0026).

The problem with pinch interaction was deciding when to let go of the plane as the release gesture was not correctly detected at a proper time as P11 mentioned that, “there was a slight delay when releasing the object which caused it to misalign several times.” This made it difficult for participants to understand when to release. P2 said, “it was hard to understand when to let go”. Participants wanted a feedback of when the plane was released as it might help them to coordinate their actions appropriately, P12 suggested, “some indication of when it is going to release soon.” Fig. 7.11 shows an example of how the accuracy changes from when the plane is placed in the final position using pinch and when the system detects the pinch to be released.

Participants faced issues in grabbing the handles in custom axis with C/D gain. P5 said that, “Sometimes the system did not recognise my pinch. I felt unsure to use it.” P4 said that, “It was hard for me to use translation handle of widget”. Fig. 7.12 shows an example of how the accuracy changes from when the plane is placed in the final position using translation handle while grabbed and being released. The accuracy change is not as high as pinch. The usability issue related to translation handle could be because of the learnability and forming a mental model as mentioned earlier. However, using rotation handle was not that difficult. P6 said that, “the most difficult part is to use the translation, but rotation is good”. P4 said that, “Rotation handle of the widget was designed natural so it was easy to



(a) Plane is grabbed and manipulated using pinch



(b) The plane is released

Figure 7.11 *An example of gesture recognition delay for pinch interaction technique*

learn and perform task with it". The usability issues with custom axis with C/D gain made the participants feel that it took them more time to finish the task, this corroborates with Fig. 7.2. P9 said that, "it took long time to perform the tasks".

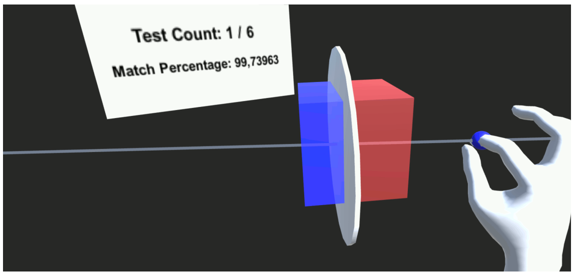
There were a couple of usability issues with push and poke, especially when participants had to push and poke with a hand behind the plane. The hand was occluded by the plane and it made it hard to decide how much effort to apply. P10 said that, "it is important that the hand does not visually cover the objects". P8 said that, "pinching on the other hand needed strong precision and more concentration". P12 said that, "it was hard to poke when error is on one side and you have to poke on other side because I was not able to see hands, occluded by plane".

7.2.3 Naturalness

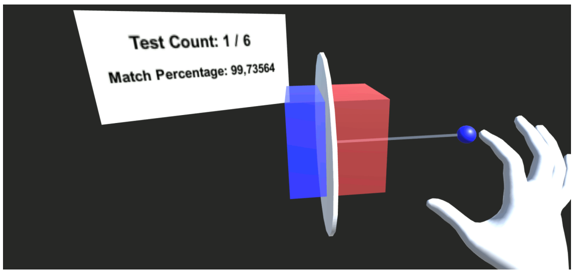
Participants felt that custom axis with C/D gain was significantly less natural than both pinch (p=0.0) and push and poke (p=0.0).

Participants felt that the motions required for push and poke were natural as P3 mentioned, "if felt organic. The movement was natural". In addition, using both hands in this technique also seemed natural. P11 said that, "using both hands to manipulate the object felt more natural and precise".

Participants felt that the usage of hands in the custom axis with C/D gain interaction was not what we use in daily life. P6 said that, "The hand movements were not natural, the method is not close to our daily hand motions. So it felt not



(a) Plane is grabbed and manipulated using pinch



(b) The plane is released

Figure 7.12 *An example of gesture recognition delay for custom axis with C/D gain interaction technique*

natural". P10 said that, "It was not so natural because I handled the plane with widgets and not with my hands". One participant felt that the hand usage did not feel natural due to the amount of hand movements. P12 said that, "When there is a lot of movement, it is hard to use it".

7.2.4 Confidence

Participants were significantly more confident with push and poke than both pinch (p=0.0024) and custom axis with C/D gain (p=0.0043).

The participants felt that they should have been able to move the plane in one interaction as, P12 mentioned, "I know how to move it, but it is hard to get it right on the first trial". The improper release functionality caused this technique to be more challenging than they expected, P8 mentioned that, "placing the plane in the right position was more challenging that i expected." The problem in the release functionality in pinch interaction made it hard to achieve the the last few percentages in accuracy, P10 mentioned that, "the last tiny corrections were difficult to make and not so accurate as I would have hoped." Participants felt a lack of confidence even though they could were handling the plane directly as P8 mentioned, "was more defective even though I had a grip of the object."

Participants had issues in grabbing the handles for custom axis with C/D gain condition. P3 said that, "It is sometimes confusing to grab red ball or blue ball as it switches automatically". The usability issues with the translation handle made it hard to make it precise. P4 said that, “It was hard for me to use translation handle of widget and that affected the task precision and time”. Some participants felt confident with this approach. P8 said that, “I was very confident in the last exercises and my aim was more precise.” P1 said that, “Custom axis + C/D Gain was hard to move but if used eventually the most correct”. In particular, those participants appreciated the rotation control. P3 said that, “two axis rotation gives much control for precise tasks.” Some participants felt it was more precise than poke interaction. P12 said that, “This has more control than poke”.

7.2.5 Precision

There was no statistical significance in terms of precision between the conditions. However, there is a statistical difference in the ranks of the precision rating across the tasks between all the conditions using Friedman Chi Square (p=0.002). The average rank of push and poke is 1.625, the average rank of custom axis with C/D gain is 1.875 and the average rank of pinch is 2.5.

When asked to rate the conditions on terms of precision, most participants felt that push and poke was more precise than custom axis with C/D gain and finally pinch as shown in Tab. 7.2.

Condition	Ranking		
	1st	2nd	3rd
Pinch	0	4	8
Push and poke	10	2	0
Custom axis with C/D gain	2	6	4

Table 7.2 *Accumulated count of users preference based on their ranking for which interaction technique felt more precise. One is the most precise and three is least precise.*

The participants found it hard to control the plane with the pinch interaction especially for minor adjustments as P5 mentioned, “small and accurate movements were difficult because if I wanted to move only one corner of the plane, it happened that the whole plane moved and changed its angle.” Participants felt that pinch was not precise due to the problem in release detection.

Participants included easiness while ranking push and poke highest for precision. P4 said that, “Push+Poke was easy to use while precising the task”. P9 said that, “The pinch and push + poke were easy to manipulate and get the object in the place”. P8 said that, “Push+Poke gave me freedom to correct the position with a small effort”.

A few participants felt that custom axis with C/D gain allowed them to make precise movements. P5 said that, “moving and rotating were separate, easy to do precise movements”. P12 said that, “Rotation in custom axis gives more control for small movements". P3 said that, "Custom axis is much precise". Learnability was a problem with other participants, however some did mention that training could have helped. P7 said that, "I think with push and poke I achieved highest precision, but if I could learn custom axis, I could probably achieve similar precision". To further understand, whether experience made any difference in subjective and objective precision, the correlation is computed. Fig. 7.14 shows the relationship between the participants' experience in using hands in virtual reality and mean accuracy achieved by them in the last three trials. The Pearson's correlation coefficient (r) between the participants' experience in using hands in virtual reality and mean accuracy achieved by them in the last three trials is 0.79. Fig. 7.13 shows the relationship between the participants' experience in using hands in virtual reality and their rating of subjective precision of custom axis with C/D gain interaction technique. The Spearman's rank correlation coefficient (ρ) between participants' experience in using hands in virtual reality and subjective precision of custom axis with C/D gain is 0.72.

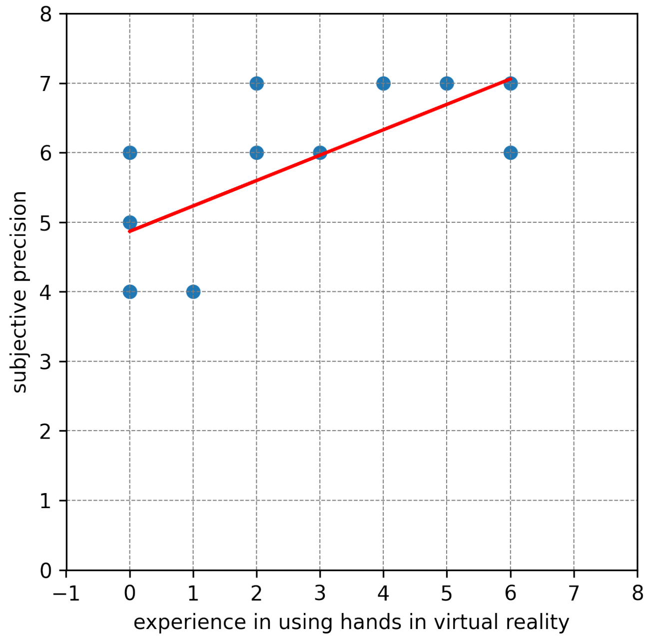


Figure 7.13 *The relationship between participants' experience in using hands in virtual reality and their rating of subjective precision of custom axis with C/D gain interaction technique.*

7.2.6 Preference

When asked to rate the conditions based on preference, most participants liked push and poke first then pinch and finally custom axis with C/D gain as mentioned in

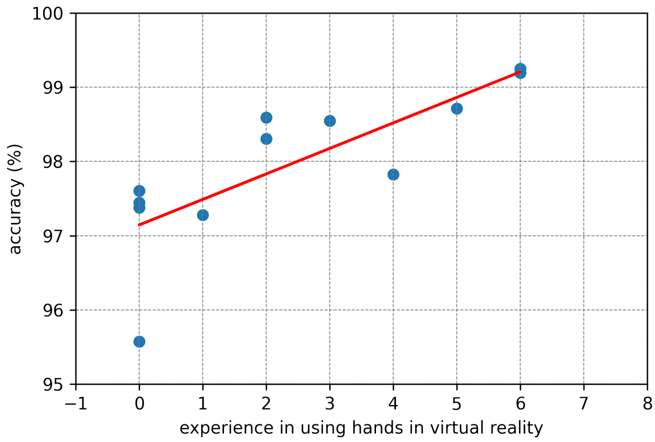


Figure 7.14 *The relationship between participants' experience in using hands in virtual reality and mean accuracy achieved by them in the last three trials of custom axis with C/D gain condition.*

Tab. 7.3.

Condition	Ranking		
	1st	2nd	3rd
Pinch	0	8	4
Push and poke	11	1	0
Custom axis with C/D gain	1	3	8

Table 7.3 *Accumulated count of users preference based on their ranking for which interaction technique they liked the most. One is the most liked and three is least liked.*

Participants considered various factors while ranking the conditions. Participants considered combinations of ease, preciseness and naturalness. P4 said that, "Push and poke was precise in performing tasks and easy to learn. Pinch is natural but precision was difficult to handle. Custom Axis + C/D Gain was difficult to learn and work with." P6 said that, "ranked based on accuracy, natural and easy of use". P10 said that, "Push and poke felt natural to use, it allowed manipulation with both hands at the same time and felt also accurate. Pinch felt okay to use but it did not feel accurate. Custom axis + C/D Gain felt accurate but it was most difficult and unnatural to use."

7.2.7 Most potential for future development

When asked to rate the conditions based on potential for future development, most participants felt that push and poke had more potential then pinch and finally custom axis with C/D gain as mentioned in Tab. 7.4.

Participants found push and poke condition fun and wanted to see how it can

Condition	Ranking		
	1st	2nd	3rd
Pinch	2	8	2
Push and poke	8	2	2
Custom axis with C/D gain	2	2	8

Table 7.4 *Accumulated count of users preference based on their ranking for which interaction technique they preferred should be developed further in future. One is for most likely to improve and three is for least likely to improve.*

be further developed. P5 said that, “I would like to see how good push+poke can be after development, because it is already good”.

Participants felt that pinch could be further developed to show feedback when it was being grabbed and released. P2 suggested “if pinch can be developed more e.g. when and how to let go”. P12 suggested, “some indication of when it is going to release soon”. A few participants wanted pinch to be made a two handed operation so that it can be made more controlled. P12 suggested “in pinch, there should be a gesture for the non-dominant hand to lock the position of the plane”. P4 suggested, “to lock other side while rotating like the rotation in custom axis with C/D gain method” and P10 suggested a similar idea, “stabilize one edge with my left hand and at the same time e.g. made a rotational movement with my right hand”.

Participants felt that the grabbing handles could be made easier. P3 mentioned, “the switching between red and blue balls can be worked in future”.

Participants also suggested combinations of techniques as these techniques could complement each other. P5, P6 and P10 suggested the combination of “pinch and poke”. P4 and P12 suggested, “pinch and rotation handle with C/D gain”. Fig. 7.15 shows an example of the trend when pinch and poke interaction techniques are combined. The first three seconds from pinch trend line are considered and the last twelve seconds of push and poke trend line are combined to create a trend for this combination.

7.3 Summary of results

Tab. 7.5 shows a summary of the results discussed above; ✓represents that there is a significant difference and interaction technique is higher in this measure than the other technique(s) marked with X and the numerical values indicate the average ranks or subjective ranks provided by users. In summary, the push and poke is faster, easy to learn and use, participants are confident using it, ranks high in subjective precision and most preferred. Custom axis is the most precise after push and poke. Pinch is second most fast, natural, easy to learn, preferred after push and poke.

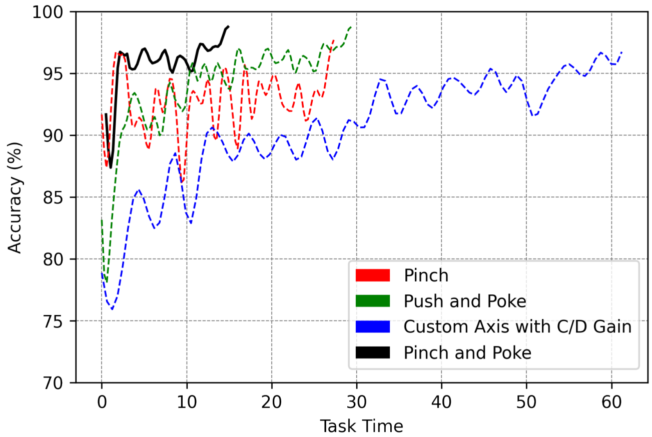


Figure 7.15 *Expected trend line when pinch and poke interaction techniques are combined.*

	Pinch	Push and poke	Custom axis with C/D gain
Task Completion Time	✓	✓	X
Execution time	X	✓	X
Evaluation time	X	✓	X
Learnability	2	1	3
Ease of use		✓	X
Naturalness	✓	✓	X
Confidence	X	✓	X
Precision	2.5	1.625	1.875
Preference	2	1	3
Potential development	2	1	3

Table 7.5 *Summary of the results of the study.*

8 Discussion

In this section, the findings and design implications from the results of the study are discussed. The limitations and future work are also discussed.

8.1 Findings

The findings from the results of the study are as follows:

8.1.1 Precise and preferred interaction technique

There is no statistical difference between the interaction techniques in terms of both objective and subjective accuracy. The Friedman Chi Square statistic showed a clear ranking of interaction techniques in terms of objective precision with push and poke being first (rank=1.25), followed by custom axis with C/D gain (rank=2.08) and finally pinch (rank=2.67). The same ranking order is seen with subjective precision: push and poke being first (rank=1.25), followed by custom axis with C/D gain (rank=1.875) and finally pinch (rank=2.5). Based on the subjective ranking and ratings, participants preferred push and poke interaction technique over both pinch and custom axis with C/D gain. This could be due to other attributes as participants felt more confident to use it, easy to learn, easy to use.

8.1.2 Experienced users perform well using tools

Most participants preferred push and poke interaction technique over custom axis with C/D Gain as push and poke was easier to learn. When asked to rank interaction techniques based on suitability for novice users, participants ranked custom axis with C/D gain the lowest. Some participants mentioned that they needed more time to practice and use custom axis. P6 said that, "need more time to practise". P7 said that, "if you get used to it, you could do it more precisely". This means that the 6 trials was not enough for certain participants to learn and use custom axis with C/D gain with enough ease. This is supported by there being no statistical difference between accuracy of the first trial and the last three trials of custom axis.

In addition, participants with experience rating of 4 to 7 in using hands in virtual reality mentioned that they could easily use custom axis with C/D gain to perform precise manipulations. P4 with experience level of 5 said that, "rotation handle of the widget was designed natural so it was easy to learn and perform task with it". P5 with experience level of 6 said that, "rotation is good" and "moving and rotating were separate so it is easy to do precise movements". P12 with experience level of 6 said that, "this has more control than poke" and "rotation in custom axis gives more control for small movements". Fig. 7.14 and Fig. 7.13 shows the relationship between the participants' experience in using hands in virtual reality and their objective and subjective precision while using custom axis with C/D gain interaction technique. The correlation for objective precision is 0.79 and subjective precision is 0.72. This indicates that more experience in using hands in virtual reality would improve user's precision in plane alignment task.

The results of the study indicate that interaction techniques such as custom axis with C/D gain could be useful for experienced users. This study did not recruit enough experienced users to statistically validate this. Further studies could be conducted to understand if there are gains in usability or time for experienced users using custom axis with C/D gain.

8.2 Design implications

The design implications based on the study are as follow.

8.2.1 Necessary that the interaction technique supports small movements

According to Graham and MacKenzie (1996), in object selection task the user movement can be divided in to two phases: (1) initial fast and imprecise movement to the target and (2) final slow and precise movements before touching the object. To understand whether this holds for object manipulation task in virtual reality, the trend of movements made are used. Since the movement can be computed in terms of both position and rotation, changes in accuracy (which uses differences in both position and orientation between precut plane and current plane) is used as an indicator of movement. Fig. 7.7 shows the overall trend lines for the three interaction techniques.

From this Fig. 7.7, it is clear that participants made small movements using push and poke as well as custom axis with C/D gain which created incremental small improvements in accuracy across time. It is also interesting to notice that the push and poke method helped the users to achieve slightly higher accuracy and this accuracy is increased much faster than custom axis with C/D gain. On the other hand, pinch did not offer these small movements due to which there were not small incremental improvements in accuracy using pinch. Instead, the first interaction caused a large movement towards a high accuracy and the later interactions oscillate around this high accuracy. This trend indicates that the participants got the plane very close to the final position but were facing issues in trying to get the plane in position and attempted several times. This is supported by participant's quotes. P11 said that "there was a slight delay when releasing the object which caused it to misalign several times". P12 said that "I know how to move it, but it is hard to get it right on the first. Required multiple trials".

Fig. 7.8 shows the distribution of the median of the minimum accuracy changes across participants for each interaction technique. The ranges of minimum movement in push and poke and custom axis with C/D gain for the participants is much less than the pinch. This difference is not significantly different but participants felt that the difference contributed in the ability of manipulating the object over smaller distance or angles.

When asked to rate and rank these interactions based on precision, participants mentioned how these interaction techniques performed in manipulating the object over smaller distance or angles. P12 said that "rotation in custom axis gives more control for small movements" while P8 said that "push and poke gave me freedom to correct the position". P10 said that "The last tiny corrections were difficult to make and not so accurate as I would have hoped". For pinch interaction, P5 said that "small and accurate movements were difficult because if I wanted to move only one corner of the plane, it happened that the whole plane moved and changed its angle". Some participants felt this could be due to the noise of hand tracking itself. P8 said that "it was difficult to place the plane precisely, probably my hand is more shaky than I imagine". P10 said that "the movement of the plane was jitterish which made it difficult to place". This is supported by participants suggestions to make pinch support small movements by making it a two handed interaction. P12 suggested "in pinch, there should be a gesture for the non-dominant hand to lock the position of the plane". P4 suggested, "to lock other side while rotating like the rotation in custom axis with C/D gain method". P10 suggested a similar idea, "stabilize one edge with my left hand and at the same time e.g. made a rotational movement with my right hand".

Therefore, it is necessary that an interaction technique for precise object manipulation should allow users to make small movements.

8.2.2 Support large movements in addition to small movements for efficient object manipulation

Though it is necessary that the interaction technique should support small movements as discussed in the earlier finding, participants thought that there is a value in large movements as well. Participants suggested combinations of techniques so that the overall interaction technique supports both small and large movements. P5, P6 and P10 suggested the combination of "pinch and poke". P4 and P12 suggested, "pinch and rotation handle with C/D gain". As discussed earlier, pinch supported much larger movements than both push and poke and custom axis with C/D gain.

Combining these complementary techniques, increases the range of movements that a user can perform. Fig. 7.15 shows an example of the trend when pinch and poke interaction techniques are combined. These suggestions are in line with (Graham & MacKenzie, 1996) in which the movement can be divided in to two phases: (1) initial fast and imprecise movement to the target and (2) final slow and precise movements before touching the object. Such combinations would be essential for reducing the interaction time as shown in Fig. 7.15, since pinch could be used to quickly reach a high precision and then the interaction technique for small movements (poke or rotation handle) could be used to increase the precision in small amounts till a good enough precision is reached.

Thus, interaction techniques could support both small and large movements to provide precision in a smaller time.

8.2.3 Designing interaction techniques which does not require gesture recognition

In addition to smaller movements, a participant mentioned that cognitive effort required to align the plane was another factor when rating the precision of interaction techniques. P8 said that, "... gave me freedom to correct the position with a small effort". In addition to effort, participants mentioned that it took them more time to use custom axis with C/D gain, even though there is no significant difference in terms of task completion time between push and poke and custom axis with C/D gain. P9 said that, "The interaction technique was very difficult to use as compared to the other methods. I took long time to perform the tasks." This perceived increase in task completion time could be attributed to cognitive effort (Cooper-Martin, 1994).

There are several ways to measure cognitive effort such as response time, decision time, secondary tasks, subjective task completion time and subjective cognitive load (Bettman et al., 1990; Cooper-Martin, 1994). The subjective cognitive load for each interaction was not collected in the study. According to the the seven stages of action proposed by Norman (2013), interaction consists of evaluation stage and execution stage. Fig. 7.3 shows the mean of the mean evaluation time in a task across the trials in different conditions and Fig. 7.4 shows the mean of the mean execution time in a task across the trials in different conditions. These figures shows that the user had to use significantly less decision making in push and poke condition than other two conditions and less execution time in push and poke than only pinch for the plane alignment task.

Participants felt that both pinch and push and poke were easier to learn due to the familiarity of the gestures used. P12 said that "pinch seemed like something I do in real life" and P5 said that, "pinch and push are familiar from real world". P4 said that, “Push and poke are daily based gestures that can be easy for a novice user to work with”. Some participants felt that the rotation handle in custom axis with C/D gain was natural to use. P4 said that, “using rotation handle was not that difficult”, P6 said that, “rotation is good” and P4 said that, “rotation handle of the widget was designed natural so it was easy to learn and perform task with it”. Familiarity has been suggested to be one of the ways for creating universal design so that a wider range of users can use the system (Turner & Van De Walle, 2006). In a study conducted by Van de Walle et al. (2003), familiarity helped the users with understanding and doing without thinking. Therefore, we expect the execution times for pinch to be less than that for custom axis with C/D gain. However, median of execution time for pinch is higher than that of custom axis with C/D gain.

As mentioned in Sec. 7.2.2, participants faced usability issues while using pinch and custom axis with C/D gain. P2 said, “it was hard to understand when to let go” for pinch. P5 said that, “Sometimes the system did not recognise my pinch. I felt unsure to use it” for custom axis with C/D gain. P4 said that, “It was hard for me to use translation handle of widget”. P12 said that, “grabbing blue ball is hard”. Both pinch and each of the handles in custom axis with C/D gain need pinch start and release gestures to be recognized. These interaction techniques allow the user to manipulate it when the grab gesture is recognized and before release gesture is recognized. Both these operations had issues related to gesture recognition specifically delay in gesture recognition as P11 mentioned “there was a slight delay when releasing the object which caused it to misalign several times.”

In the case of pinch, if the gesture recognition is delayed, the fingers are in a different unrelated position and then plane’s position and orientation are adjusted to the fingers’ positions at that delayed time as shown in Fig. 7.11. In addition, noisy hand tracking could contribute to a noisy fingers’ positions to be detected. This is the reason behind the pinch interaction technique not being support small enough movements and not being that precise. Custom axis with C/D gain uses C/D gain to scale down the amount of movement. When there is a delay in gesture recognition and noisy hand tracking, the error in plane movement is constrained along with axis and reduces down by C/D gain as shown in Fig. 7.12. Thus in the case of custom axis with C/D gain, constraining the motion along an axis and scaling down the movements with C/D gain compensate for the delay in gesture recognition and noisy hand tracking.

Therefore, it is better to design interaction techniques without any gesture recognition. If it is not possible to design interaction techniques without gesture recognition, then having strategies to compensate for gesture recognition delays and noisy hand tracking is the next best option.

8.3 Limitations

During the study the users were asked to take their time to practise and learn the custom axis with C/D gain interaction method. Still some of the user felt they required more time to practise that method to learn and use the interaction method effectively. The results of the study may differ if more time was provided to participants to practice this interaction technique. In addition, more participants with high levels of experience in using hands in virtual reality could be recruited to find an interaction technique well suited for experienced users.

During this study, we did not collect any subjective cognitive load from the user to measure the cognitive effort. Instead of calculating the evaluation and execution time and using it to infer cognitive load, subjective load could have been collected from participants.

Even though this study is about object manipulation, this study is designed for the context of osteotomy operation planning due to which the object that we manipulated throughout this study was a plane. Though these interaction techniques can be generalised to other 3D objects in case of custom axis and C/D gain modifications are necessary to work with other 3D objects.

In case of physics based push and poke interaction technique, the physics settings of handling the plane which is mentioned in Sec.5.2.2 were fixed based on the plane alignment task and initial pilots. Due to these fixed setup the users were suggested not to move the plane at higher speed for Custom axis with C/D gain based method. The settings of this interaction could have been explored to support a larger range of movement and there could have been changes in the accuracy trend and the final precision achieved with the interaction technique.

8.4 Future Work

Future work would explore designing interaction techniques that support both small and large movements. Possible interaction techniques that could be designed include (1) two handed pinch, (2) pinch and poke, (3) pinch and rotation handle of custom axis with C/D gain. In addition, some form of feedback, either haptic, auditory or visual could be provided for indicating states such as when plane is touched, interaction has started and stopped. A controlled study could be performed to understand which interaction technique is most precise, efficient and preferred.

9 Conclusion

This research work designed and evaluated precise hand-based interaction techniques for plane alignment task in virtual reality. To answer the research questions, a controlled study was conducted in which the interaction techniques were used in counterbalanced order and plane alignment was used as the task.

The answers to these research questions are:

1. Push and poke interaction technique is subjectively ranked more precise and preferred.

Push and poke interaction technique is subjectively ranked more precise and preferred because it is easy to learn, easy to use and participants felt more confident while using it. The results of the study indicate that interaction techniques such as custom axis with C/D gain could be useful for experienced users. Further studies could be conducted to understand if there are gains in usability or time for experienced users using custom axis with C/D gain.

2. Interaction techniques for precise object manipulation should support smaller movements. Interaction techniques could support large movements in addition to small movements for efficient object manipulation.

The trend lines and participant quotes show that interaction techniques for precise object manipulation should be able to make small movements. Participants made small movements using push and poke as well as custom axis with C/D gain which created incremental small improvements in accuracy across time. On the other hand, pinch did not offer these small movements due to which there were not small incremental improvements in accuracy using pinch. Instead, the first interaction caused a large movement towards a high accuracy and the later interactions oscillate around this high accuracy. When asked to rate and rank these interactions based on precision, participants ranked interaction techniques that provided small movements higher than those that support large movements. However, participants also suggested combinations of techniques so that the overall interaction technique supports both small and large movements. Such combinations would be essential for reducing the interaction time since pinch could be used to quickly reach a high precision and then the interaction technique for small movements (poke or rotation handle) could be used to increase the precision in small amounts till a good enough precision is reached.

3. Interaction techniques should try to avoid gesture recognition and if not then strategies to compensate for the delay in gesture recognition

and noisy hand tracking should be incorporated.

Pinch and custom axis with C/D gain interaction techniques increased the cognitive load of the user. Participants faced usability issues with these interaction techniques due to delay in gesture recognition and noisy hand tracking. In the case of pinch, if the gesture recognition is delayed, the plane's position and orientation are adjusted to the fingers' noisy positions at that delayed time. On the other hand, when there is a delay in gesture recognition and noisy hand tracking in custom axis with C/D gain, the error in plane movement is constrained along with axis and reduces down by C/D gain. The custom axis and C/D gain design factors act as strategies to compensate for the delay in gesture recognition and noisy hand tracking. Thus, it is better that gesture recognition is not used and if it has to be used then strategies to compensate have to be used.

The contributions of this study are (1) taxonomy of design factors for hand interaction techniques for object manipulation in virtual reality, (2) proposed designs for interaction techniques for plane adjustment task for performing jaw osteotomy operation planning in virtual reality, (3) empirical validation of interaction techniques for plane alignment task, (4) design implications for future hand based interaction techniques for precise plane alignment in virtual reality.

The interaction techniques in this research work have been designed for plane alignment task. These interaction techniques could be used for medical operating planning steps which requires 3D objects to be placed precisely. Push and poke could be used to push and poke the boundaries of the 3D object. The handles of custom axis with C/D gain can be created on the closest surface of the 3D object. Future studies could be performed to understand how these interaction techniques perform in specific medical planning contexts.

Using controllers as an input device in virtual reality has been the normal till now. But controllers have limited input space and our hands are always required to hold it. It is also required to be charged and so it is not always available. On the other side, hands are always available, and have larger input space compared to a controller. In the real world most of the times we use our hands to interact. Our hands are versatile and would require no or less learning for any user to get started and use as an input medium in virtual reality. Using hand as an input for interacting in virtual reality is not reliable for all scenarios currently and users currently find using controllers for interacting in virtual reality to be more accurate and reliable (Caggianese et al., 2018; Galais et al., 2019; Gusai et al., 2017). Earlier studies have revealed users like using their hands as an input device in virtual reality and augmented reality (Figueiredo et al., 2018). So it is crucial to identify, design and develop different interaction and user interface elements that are suited for hand based interaction.

This study has shown that it is possible to design interaction techniques that do not use gesture recognition and still be precise than the baseline methods for precise object manipulation in virtual reality. Interaction techniques, that do not use gesture recognition, do not face the usability issues created by delay in gesture recognition. In the future, interactions techniques that do not use gesture recognition should be explored for virtual reality applications.

Bibliography

Accot, J., & Zhai, S. (2003). Refining fitts' law models for bivariate pointing. *Proceedings of the SIGCHI conference on Human factors in computing systems*, 193–200.

Ahmad, S., & Tresp, V. (1993). Classification with missing and uncertain inputs. *IEEE International Conference on Neural Networks*, 1949–1954.

Arora, R., Kazi, R. H., Kaufman, D. M., Li, W., & Singh, K. (2019). Magicalhands: Mid-air hand gestures for animating in vr. *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, 463–477.

Bettman, J. R., Johnson, E. J., & Payne, J. W. (1990). A componential analysis of cognitive effort in choice. *Organizational behavior and human decision processes*, *45*(1), 111–139.

Blackley, S. V., Huynh, J., Wang, L., Korach, Z., & Zhou, L. (2019). Speech recognition for clinical documentation from 1990 to 2018: A systematic review. *Journal of the american medical informatics association*, *26*(4), 324–338.

Boléo-Tomé, J. (1998). *Aesthetic plastic surgery*, *22*(3), 185–189.

Bossavit, B., Marzo, A., Ardaiz, O., De Cerio, L. D., & Pina, A. (2014). Design choices and their implications for 3d mid-air manipulation techniques. *Presence: Teleoperators and Virtual Environments*, *23*(4), 377–392.

Boukhayma, A., Bem, R. d., & Torr, P. H. (2019). 3d hand shape and pose from images in the wild. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 10843–10852.

Bowman, D. A., & Hodges, L. F. (1999). Formalizing the design, evaluation, and application of interaction techniques for immersive virtual environments. *Journal of Visual Languages & Computing*, *10*(1), 37–53.

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

Brewster, L. J., Trivedi, S. S., Tuy, H. K., & Udupa, J. K. (1984). Interactive surgical planning. *IEEE Computer Graphics and Applications*, *4*(3), 31–40.

Caggianese, G., Gallo, L., & Neroni, P. (2018). The vive controllers vs. leap motion for interactions in virtual environments: A comparative evaluation. *International Conference on Intelligent Interactive Multimedia Systems and Services*, 24–33.

Canales, R., & Jörg, S. (2020). Performance is not everything: Audio feedback preferred over visual feedback for grasping task in virtual reality. *Motion, interaction and games* (pp. 1–6).

Cao, Z., Hidalgo, G., Simon, T., Wei, S.-E., & Sheikh, Y. (2019). Openpose: Realtime multi-person 2d pose estimation using part affinity fields. *IEEE transactions on pattern analysis and machine intelligence*, *43*(1), 172–186.

Caputo, F. M., Emporio, M., & Giachetti, A. (2018). The smart pin: An effective tool for object manipulation in immersive virtual reality environments. *Computers & Graphics*, *74*, 225–233.

Cho, I., & Wartell, Z. (2015). Evaluation of a bimanual simultaneous 7dof interaction technique in virtual environments. *2015 IEEE symposium on 3D User Interfaces (3DUI)*, 133–136.

Clohisy, J. C., Schutz, A. L., John, L. S., Schoenecker, P. L., & Wright, R. W. (2009). Periacetabular osteotomy: A systematic literature review. *Clinical Orthopaedics and Related Research®*, *467*(8), 2041–2052.

Cooper-Martin, E. (1994). Measures of cognitive effort. *Marketing Letters*, *5*(1), 43–56.

Cootes, T. F., Taylor, C. J., Cooper, D. H., & Graham, J. (1995). Active shape models-their training and application. *Computer vision and image understanding*, *61*(1), 38–59.

Corbett, B., Nam, C. S., & Yamaguchi, T. (2016). The effects of haptic feedback and visual distraction on pointing task performance. *International Journal of Human-Computer Interaction*, *32*(2), 89–102.

Cruz, A., & Green, B. G. (2000). Thermal stimulation of taste. *Nature*, *403*(6772), 889–892.

Cruz Cebrian, A. (2017). Enhancing vr experiences with blowing input techniques [Accessed: 2020-04-04].

Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., Kenyon, R. V., & Hart, J. C. (1992). The cave: Audio visual experience automatic virtual environment. *Communications of the ACM*, *35*(6), 64–73.

Cutler, L. D., Fröhlich, B., & Hanrahan, P. (1997). Two-handed direct manipulation on the responsive workbench. *Proceedings of the 1997 symposium on Interactive 3D graphics*, 107–114.

Di Matteo, B., Tarabella, V., Filardo, G., Viganò, A., Tomba, P., & Marcacci, M. (2013). John rhea barton: The birth of osteotomy. *Knee Surgery, Sports Traumatology, Arthroscopy*, *21*(9), 1957–1962.

Dmitrenko, D., Maggioni, E., & Obrist, M. (2017). Ospace: Towards a systematic exploration of olfactory interaction spaces. *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces*, 171–180.

Ellis, S. R. (1994). What are virtual environments? *IEEE Computer Graphics and Applications*, *14*(1), 17–22.

Erickson, T. D. (1995). Working with interface metaphors. *Readings in human–computer interaction* (pp. 147–151). Elsevier.

Feix, T., Romero, J., Schmiedmayer, H.-B., Dollar, A. M., & Kragic, D. (2015). The grasp taxonomy of human grasp types. *IEEE Transactions on human-machine systems*, *46*(1), 66–77.

Figueiredo, L., Rodrigues, E., Teixeira, J., & Teichrieb, V. (2018). A comparative evaluation of direct hand and wand interactions on consumer devices. *Computers & Graphics*, *77*, 108–121.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology*, *47*(6), 381.

Fitzmaurice, G. W., Zhai, S., & Chignell, M. H. (1993). Virtual reality for palmtop computers. *ACM Transactions on Information Systems (TOIS)*, *11*(3), 197–218.

Freeman, W. T., & Roth, M. (1995). Orientation histograms for hand gesture recognition. *International workshop on automatic face and gesture recognition*, *12*, 296–301.

Frees, S., & Kessler, G. D. (2005). Precise and rapid interaction through scaled manipulation in immersive virtual environments. *IEEE Proceedings. VR 2005. Virtual Reality, 2005.*, 99–106.

Galais, T., Delmas, A., & Alonso, R. (2019). Natural interaction in virtual reality: Impact on the cognitive load. *Proceedings of the 31st Conference on l'Interaction Homme-Machine: Adjunct*, 1–9.

Ge, L., Ren, Z., Li, Y., Xue, Z., Wang, Y., Cai, J., & Yuan, J. (2019). 3d hand shape and pose estimation from a single rgb image. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 10833–10842.

Graham, E. D., & MacKenzie, C. L. (1996). Physical versus virtual pointing. *Proceedings of the SIGCHI conference on Human factors in computing systems*, 292–299.

Guiard, Y. (1987). Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. *Journal of motor behavior*, *19*(4), 486–517.

Gusai, E., Bassano, C., Solari, F., & Chessa, M. (2017). Interaction in an immersive collaborative virtual reality environment: A comparison between leap motion and htc controllers. *International Conference on Image Analysis and Processing*, 290–300.

Hermann, M., & Weber, M. (2009). When three worlds collide: A model of the tangible interaction process. *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7*, 341–344.

Hsieh, M.-S., Tsai, M.-D., & Chang, W.-C. (2002). Virtual reality simulator for osteotomy and fusion involving the musculoskeletal system. *Computerized medical imaging and graphics*, *26*(2), 91–101.

Jacoby, R. H., Ferneau, M., & Humphries, J. (1994). Gestural interaction in a virtual environment. *Stereoscopic Displays and Virtual Reality Systems*, *2177*, 355–364.

Kim, H., & Choi, Y. (2019). Performance comparison of user interface devices for controlling mining software in virtual reality environments. *Applied Sciences*, *9*(13), 2584.

Klatzky, R. L., Pellegrino, J., McCloskey, B. P., & Lederman, S. J. (1993). Cognitive representations of functional interactions with objects. *Memory & Cognition*, *21*(3), 294–303.

Kruger, R., Carpendale, S., Scott, S. D., & Tang, A. (2005). Fluid integration of rotation and translation. *Proceedings of the SIGCHI conference on Human factors in computing systems*, 601–610.

Lanitis, A., Taylor, C. J., Cootes, T., & Ahmed, T. (1995). Automatic interpretation of human faces and hand gestures using flexible models. *In International Workshop on Automatic Face-and Gesture-Recognition*.

Leap Motion, S. t. m. (2012). Leap motion: Introducing the skeletal tracking model [Accessed: 2020-04-04].

Lingard, B. (1995). Human interfacing issues of virtual reality [Accessed: 2020-04-04].

Mateo, J. C., Cowgill Jr, J. L., Moore Jr, T. J., Gilkey Jr, R. H., Simpson Jr, B. D., Weisenberger Jr, J. M., & Manning Jr, J. T. (2005). Evaluation of a collaborative movement task in a distributed three-dimensional virtual environment. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *49*(17), 1578–1582.

Mendes, D., Relvas, F., Ferreira, A., & Jorge, J. (2016). The benefits of dof separation in mid-air 3d object manipulation. *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*, 261–268.

Mendes, D., Sousa, M., Lorena, R., Ferreira, A., & Jorge, J. (2017). Using custom transformation axes for mid-air manipulation of 3d virtual objects. *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, 1–8.

Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1995). Augmented reality: A class of displays on the reality-virtuality continuum. *Telemanipulator and telepresence technologies*, *2351*, 282–292.

Mine, M. R., Brooks Jr, F. P., & Sequin, C. H. (1997). Moving objects in space: Exploiting proprioception in virtual-environment interaction. *Proceedings of* *the 24th annual conference on Computer graphics and interactive techniques*, 19–26.

Mlyniec, P., Jerald, J., Yoganandan, A., Seagull, F. J., Toledo, F., & Schultheis, U. (2011). Imedic: A two-handed immersive medical environment for distributed interactive consultation. *MMVR*, 372–378.

Murata, A., & Iwase, H. (2001). Extending fitts' law to a three-dimensional pointing task. *Human movement science*, *20*(6), 791–805.

Nakamura, H., & Miyashita, H. (2011). Augmented gustation using electricity. *Proceedings of the 2nd augmented human international conference*, 1–2.

Nguyen, T. T. H., Duval, T., & Pontonnier, C. (2014). A new direct manipulation technique for immersive 3d virtual environments. *ICAT-EGVE 2014: the 24th International Conference on Artificial Reality and Telexistence and the 19th Eurographics Symposium on Virtual Environments*, 8.

Norman, D. (2013). *The design of everyday things: Revised and expanded edition.* Basic books.

Oberweger, M., & Lepetit, V. (2017). Deepprior++: Improving fast and accurate 3d hand pose estimation. *Proceedings of the IEEE international conference on computer vision Workshops*, 585–594.

Oculus Quest, H. T. (2020). Oculus for developers: Enable hand tracking [Accessed: 2020-04-04].

Oikonomidis, I., Kyriazis, N., & Argyros, A. A. (2011). Efficient model-based 3d tracking of hand articulations using kinect. *BmVC*, *1*(2), 3.

Olsson, P., Nysjö, F., Rodríguez-Lorenzo, A., Thor, A., Hirsch, J.-M., & Carlbom, I. B. (2015). Haptics-assisted virtual planning of bone, soft tissue, and vessels in fibula osteocutaneous free flaps. *Plastic and Reconstructive Surgery Global Open*, *3*(8).

Osawa, N. (2008). Two-handed and one-handed techniques for precise and efficient manipulation in immersive virtual environments. *International Symposium on Visual Computing*, 987–997.

Panchal-Kildare, S., & Malone, K. (2013). Skeletal anatomy of the hand. *Hand clinics*, *29*(4), 459–471.

Pausch, R., Shackelford, M. A., & Proffitt, D. (1993). A user study comparing head-mounted and stationary displays. *Proceedings of 1993 ieee research properties in virtual reality symposium*, 41–45.

Pfeuffer, K., Mayer, B., Mardanbegi, D., & Gellersen, H. (2017). Gaze+ pinch interaction in virtual reality. *Proceedings of the 5th Symposium on Spatial User Interaction*, 99–108.

Pham, D.-M., & Stuerzlinger, W. (2019). Is the pen mightier than the controller? a comparison of input devices for selection in virtual and augmented reality. *25th ACM Symposium on Virtual Reality Software and Technology*, 1–11.

Posnick, J. C. (2013). *Principles and practice of orthognathic surgery*. Elsevier Health Sciences.

Poupyrev, I., Billinghurst, M., Weghorst, S., & Ichikawa, T. (1996). The go-go interaction technique: Non-linear mapping for direct manipulation in vr. *Proceedings of the 9th annual ACM symposium on User interface software and technology*, 79–80.

Rekimoto, J., & Nagao, K. (1995). The world through the computer: Computer augmented interaction with real world environments. *Proceedings of the 8th annual ACM symposium on User interface and software technology*, 29–36.

Ruan, S., Wobbrock, J. O., Liou, K., Ng, A., & Landay, J. A. (2018). Comparing speech and keyboard text entry for short messages in two languages on touchscreen phones. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *1*(4), 1–23.

Sand, A., Rakkolainen, I., Isokoski, P., Kangas, J., Raisamo, R., & Palovuori, K. (2015). Head-mounted display with mid-air tactile feedback. *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology*, 51–58.

Santos, B. S., Dias, P., Pimentel, A., Baggerman, J.-W., Ferreira, C., Silva, S., & Madeira, J. (2009). Head-mounted display versus desktop for 3d navigation in virtual reality: A user study. *Multimedia tools and applications*, *41*(1), 161–181.

Schlenzig, J., Hunter, E., & Jain, R. (1994). Recursive identification of gesture inputs using hidden markov models. *Proceedings of 1994 IEEE Workshop on Applications of Computer Vision*, 187–194.

Schwarz, R. J., & Taylor, C. (1955). The anatomy and mechanics of the human hand. *Artificial limbs*, *2*(2), 22–35.

Sharp, T., Keskin, C., Robertson, D., Taylor, J., Shotton, J., Kim, D., Rhemann, C., Leichter, I., Vinnikov, A., Wei, Y., et al. (2015). Accurate, robust, and flexible real-time hand tracking. *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, 3633–3642.

Shigeishi, H., Ohta, K., & Takechi, M. (2015). Risk factors for postoperative complications following oral surgery. *Journal of Applied Oral Science*, *23*(4), 419–423.

Shneiderman, B. (1981). Direct manipulation: A step beyond programming languages. *Proceedings of the Joint Conference on Easier and More Productive Use of Computer Systems.(Part-II): Human Interface and the User Interface-Volume 1981*, 143.

Singh, A. K., Gramann, K., Chen, H.-T., & Lin, C.-T. (2021). The impact of hand movement velocity on cognitive conflict processing in a 3d object selection task in virtual reality. *NeuroImage*, *226*, 117578.

Slater, M. (2003). A note on presence terminology. *Presence connect*, *3*(3), 1–5.

Song, P., Goh, W. B., Hutama, W., Fu, C.-W., & Liu, X. (2012). A handle bar metaphor for virtual object manipulation with mid-air interaction. *Proceedings of the SIGCHI conference on human factors in computing systems*, 1297–1306.

Sra, M., Xu, X., & Maes, P. (2018). Breathvr: Leveraging breathing as a directly controlled interface for virtual reality games. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–12.

Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of communication*, *42*(4), 73–93.

Strickland, J. (2007). How virtual reality works. *How Stuff Works*, *29*.

Sus: A “quick and dirty’usability. (1996). *Usability evaluation in industry*, *189*.

Sutherland, I. (1965). The ultimate display.

Sutherland, I. (1968). A head-mounted three dimensional display. *Proceedings of the December 9-11, 1968, fall joint computer conference, part I*, 757–764.

Suzuki, C., Narumi, T., Tanikawa, T., & Hirose, M. (2014). Affecting tumbler: Affecting our flavor perception with thermal feedback. *Proceedings of the 11th conference on advances in computer entertainment technology*, 1–10.

Turner, P., & Van De Walle, G. (2006). Familiarity as a basis of universal design. *Journal of Gerontechnology*, *5*(3), 150–159.

Van de Walle, G., Turner, P., & Davenport, E. (2003). A study of familiarity. *Human-Computer Interaction-INTERACT*, *3*, 463–70.

Vosinakis, S., & Koutsabasis, P. (2018). Evaluation of visual feedback techniques for virtual grasping with bare hands using leap motion and oculus rift. *Virtual Reality*, *22*(1), 47–62.

Wagner, A., Rasse, M., Millesi, W., & Ewers, R. (1997). Virtual reality for orthognathic surgery: The augmented reality environment concept. *Journal of oral and maxillofacial surgery*, *55*(5), 456–462.

Wan, C., Probst, T., Van Gool, L., & Yao, A. (2018). Dense 3d regression for hand pose estimation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 5147–5156.

Wang, R., Paris, S., & Popović, J. (2011). 6d hands: Markerless hand-tracking for computer aided design. *Proceedings of the 24th annual ACM symposium on User interface software and technology*, 549–558.

Wang, Y., & MacKenzie, C. L. (1999). Object manipulation in virtual environments: Relative size matters. *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, 48–55.

Ware, C., Arthur, K., & Booth, K. S. (1993). Fish tank virtual reality. *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*, 37–42.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence*, *7*(3), 225–240.

Xia, J., Ip, H. H., Samman, N., Wang, D., Kot, C. S., Yeung, R. W., & Tideman, H. (2000). Computer-assisted three-dimensional surgical planning and simulation: 3d virtual osteotomy. *International journal of oral and maxillofacial surgery*, *29*(1), 11–17.

Xu, W., Chatterjee, A., Zollhoefer, M., Rhodin, H., Fua, P., Seidel, H.-P., & Theobalt, C. (2019). Mo 2 cap 2: Real-time mobile 3d motion capture with a cap-mounted fisheye camera. *IEEE transactions on visualization and computer graphics*, *25*(5), 2093–2101.

Zhai, S., Buxton, W., & Milgram, P. (1994). The "silk cursor" investigating transparency for 3d target acquisition. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 459–464.

Zimmermann, C., & Brox, T. (2017). Learning to estimate 3d hand pose from single rgb images. *Proceedings of the IEEE international conference on computer vision*, 4903–4911.

Appendix A: Background Questionnaire

Age: _____

Gender: [ ] Female [ ] Male [ ] Other

Occupation:
[ ] Student
[ ] Bachelor [ ] Master [ ] PhD
[ ] Teacher / lecturer / professor
[ ] Full time employee
[ ] Other: _____

Handedness: [ ] Left Handed [ ] Right Handed

What is your experience level in using virtual reality?
(1= low, 7= high)
0 1 2 3 4 5 6 7

What is your experience level with using hands in virtual reality?
(Using bare hands instead of controllers e.g. Oculus Quest)
(1= low, 7= high)
0 1 2 3 4 5 6 7

Appendix B: Subjective Condition Evaluation Questionnaire

B.1 Interaction method: pinch

Evaluate the following statements:

	Strongly disagree					Strongly agree
You were confident to use the interaction method.	1	2	3	4	5	6	7
It was natural to perform the given tasks with this interaction method.	1	2	3	4	5	6	7
It was intuitive to understand and use the interaction method.	1	2	3	4	5	6	7
It was easy to perform the given tasks with this interaction method.	1	2	3	4	5	6	7
I was able to learn to use this system quickly.	1	2	3	4	5	6	7
The plane placement was precise with this interaction method.	1	2	3	4	5	6	7
Your hands are NOT tired.	1	2	3	4	5	6	7
I can imagine using this interaction method on daily basis.	1	2	3	4	5	6	7

If you specified 4 or below for any of the above rows, what is the reason behind it?

What was positive or negative on this interaction method? Other comments related to this interaction method:

B.2 Interaction method: push and poke

Evaluate the following statements:

	Strongly disagree					Strongly agree
You were confident to use the interaction method.	1	2	3	4	5	6	7
It was natural to perform the given tasks with this interaction method.	1	2	3	4	5	6	7
It was intuitive to understand and use the interaction method.	1	2	3	4	5	6	7
It was easy to perform the given tasks with this interaction method.	1	2	3	4	5	6	7
It was easy to use the push interaction technique.	1	2	3	4	5	6	7
It was easy to use the poke interaction technique.	1	2	3	4	5	6	7
I was able to learn to use this system quickly.	1	2	3	4	5	6	7
The plane placement was precise with this interaction method.	1	2	3	4	5	6	7
Your hands are NOT tired.	1	2	3	4	5	6	7
I can imagine using this interaction method on daily basis.	1	2	3	4	5	6	7

If you specified 4 or below for any of the above rows, what is the reason behind it?

What was positive or negative on this interaction method? Other comments related to this interaction method:

B.3 Interaction method: custom axis with C/D gain

Evaluate the following statements:

	Strongly disagree					Strongly agree
You were confident to use the interaction method.	1	2	3	4	5	6	7
It was natural to perform the given tasks with this interaction method.	1	2	3	4	5	6	7
It was intuitive to understand and use the interaction method.	1	2	3	4	5	6	7
It was easy to perform the given tasks with this interaction method.	1	2	3	4	5	6	7
It was easy to use the translation (blue) handle of the widget.	1	2	3	4	5	6	7
It was easy to use the rotation (red) handle of the widget.	1	2	3	4	5	6	7
I was able to learn to use this system quickly.	1	2	3	4	5	6	7
The plane placement was precise with this interaction method.	1	2	3	4	5	6	7
Your hands are NOT tired.	1	2	3	4	5	6	7
I can imagine using this interaction method on daily basis.	1	2	3	4	5	6	7

If you specified 4 or below for any of the above rows, what is the reason behind it?

What was positive or negative on this interaction method? Other comments related to this interaction method:

Appendix C: Subjective Post Experiment Questionnaire

Rank the systems based on which you liked from best to worst:
(1 = the best, 3 = the worst)
_____ pinch _____ push+poke _____ custom axis+C/D gain
The reasons behind the ranking:

Rank the systems based on interaction technique for precise manipulation:
(1 = the best, 3 = the worst)
_____ pinch _____ push+poke _____ custom axis+C/D gain
The reasons behind the ranking:

Rank the systems based on which is best suited for a novice user:
(1 = the best, 3 = the worst)
_____ pinch _____ push+poke _____ custom axis+C/D gain
The reasons behind the ranking:

Rank the systems based on which is the most potential to be developed further in the future: (1 = the best, 3 = the worst)
_____ pinch _____ push+poke _____ custom axis+C/D gain
The reasons behind the ranking: